

## Machine-learning calibration of intense x-ray free-electron-laser pulses using Bayesian optimization

Niels Breckwoldt <sup>1,2,3,\*</sup> Sang-Kil Son (손상길) <sup>1,†</sup> Tommaso Mazza,<sup>4</sup> Aljoscha Rörig <sup>2,4</sup> Rebecca Boll <sup>4</sup>  
Michael Meyer <sup>4</sup> Aaron C. LaForge <sup>5</sup> Debadarshini Mishra <sup>5</sup> Nora Berrah <sup>5</sup> and Robin Santra <sup>1,2,3</sup>

<sup>1</sup>Center for Free-Electron Laser Science CFEL, Deutsches Elektronen-Synchrotron DESY, Notkestrasse 85, 22607 Hamburg, Germany

<sup>2</sup>Department of Physics, Universität Hamburg, Notkestrasse 9, 20607 Hamburg, Germany

<sup>3</sup>The Hamburg Centre for Ultrafast Imaging, Luruper Chaussee 149, 22761 Hamburg, Germany

<sup>4</sup>European XFEL, 22869 Schenefeld, Germany

<sup>5</sup>Department of Physics, University of Connecticut, Storrs, Connecticut 06269, USA



(Received 23 December 2022; accepted 13 March 2023; published 22 May 2023)

X-ray free-electron lasers (XFELs) have brought new ways to probe and manipulate atomic and molecular dynamics with unprecedented spatial and temporal resolutions. A quantitative comparison of experimental results with their simulated theoretical counterpart, however, generally requires a precise characterization of the spatial and temporal x-ray pulse profile, providing a nonuniform photon distribution. The determination of the pulse profile constitutes a major, yet inevitable, challenge. Here, we propose a calibration scheme for intense XFEL pulses utilizing a set of experimental charge-state distributions of light noble gas atoms at a series of pulse energies in combination with first-principles simulations of the underlying atomic x-ray multiphoton ionization dynamics. The calibration builds on Bayesian optimization, which is a powerful, machine-learning-based tool particularly well suited for computationally expensive numerical optimization. We demonstrate the presented scheme to calibrate the pulse duration as well as the spatial fluence distribution profile of XFEL pulses. Our proposed method can serve as a comprehensive tool for characterizing ultraintense and ultrafast x-ray pulses.

DOI: [10.1103/PhysRevResearch.5.023114](https://doi.org/10.1103/PhysRevResearch.5.023114)

## I. INTRODUCTION

The development of x-ray free-electron lasers (XFELs) as radiation sources of ultrashort x-ray pulses with unprecedented intensity and brilliance [1,2] has opened new pathways to probing and controlling atomic and molecular dynamics with extraordinary resolution, and consequently it has become well established for studying a huge variety of phenomena covering structural biology [3–5], ultrafast x-ray atomic and molecular physics [6–9], as well as dense matter physics [10]. Prominent examples demonstrating the outstanding opportunities of XFELs include serial femtosecond crystallography [11] and single-particle imaging [12,13], both aiming for atomic spatial and femtosecond time resolutions [14–18].

The interaction between the XFEL pulse and the target, such as molecules, atoms, and ions, is highly dependent on the photon distribution within the x-ray pulse, particularly when multiple photons are involved in the physical processes. However, the photon distribution generally does not follow a uniform but rather a nonuniform distribution depending on time, i.e., the pulse duration, as well as on the spatial pulse profile in the focal spot [19]. Therefore, a range of different photon fluence values covered by the

distribution contributes to the light-matter interaction, and simulating such experiments necessarily involves the summation of all fluence-dependent contributions calculated with the actual photon fluence. This procedure, referred to as *volume integration* [20,21], requires a precise spatial and temporal characterization of the photon distribution of the XFEL pulse.

Direct measurements of XFEL pulse fluence distributions constitute a significant experimental challenge yet to be solved. The spatial properties of a focused x-ray beam can, for instance, be characterized using the ablation imprint method [22–26]. Yet, this method cannot be employed *in situ*, impeding its inclusion into experimental routines. The Hartmann wavefront sensing method [27–30], in contrast, provides an *in situ* single-shot characterization of the spatial beam profile, but its applicability to high-fluence x-ray beams is limited. Additional methods for beam characterization include grating interferometry [31,32], the Ronchi test [33,34], curved grating monitors [35,36], diffraction with aerosol spheres [37], ptychographic imaging [38], and speckle visibility spectroscopy [39]. Although they provide the beam shape and thus the shape of the spatial fluence distribution, they generally do not retrieve the fluence values on an absolute scale required for volume integration. Additionally, second-order intensity correlation measurements, also referred to as Hanbury Brown–Twiss (HBT) interferometry, are able to explore intensity correlations at different spatial or temporal positions and may subsequently be used for determining spatial coherence and the average pulse duration [40–47].

Here, we follow a different approach: Instead of a fully experimental pulse characterization, we show that the spatial and temporal fluence distribution can be obtained from

\*Corresponding author: niels.breckwoldt@cfel.de

†Corresponding author: sangkil.son@cfel.de

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

a quantitative comparison between theoretical and experimental charge-state distributions (CSDs) that are generated by x-ray multiphoton multiple ionization of light noble gas atoms, e.g., neon (Ne), resulting from the interaction with ultraintense and ultrashort XFEL pulses. The exposure to such strong x-ray pulses results in x-ray multiphoton ionization [48] characterized by a sequence of photoabsorption events accompanied by Auger-Meitner and x-ray fluorescence decays [49]. Multiple photoionizations along with subsequent decay cascades can lead to highly charged states of the involved atoms [20,50,51] or molecules [52]. Measurements of generated ion fragments are relatively straightforward, only requiring ion time-of-flight (TOF) detection, which is accessible at nearly every XFEL instrument dedicated to gas-phase studies [53–55]. As the observed ion yields represent ensemble-averaged quantities accumulated over many shots, our aim is an ensemble-averaged rather than a single-shot pulse characterization.

The obtained experimental results are subsequently compared to their numerically simulated equivalents. For the latter, we employ the *ab initio* electronic-structure toolkit XATOM [56,57], which, for any given element and electronic configuration, calculates orbitals, orbital energies, photoionization cross sections, and fluorescence and Auger-Meitner rates based on a Hartree-Fock-Slater approach. The x-ray multiphoton ionization dynamics are then simulated by solving a set of coupled rate equations employing the calculated rates and cross sections. This approach has been well tested with a series of gas-phase XFEL experiments (for a review, see Ref. [57] and references therein). As demonstrated in Ref. [21], using CSDs of (light) atoms, like Ne or argon (Ar), comes with several advantages as they are highly sensitive to the peak fluence as well as the fluence profile in the focal spot due to the high nonlinearity of the XFEL-matter interaction. Furthermore, the number of involved electronic configurations of light atoms compared to heavier atoms is small, thus reducing the computational effort for simulating the ionization dynamics.

For a quantitative comparison with the experimental results, the CSDs obtained from XATOM simulations have to be volume-integrated. To do so, we assume a specific functional form of the spatial and temporal fluence profile at the focal spot, depending on a few parameters, instead of a pointwise mapping of the fluence distribution as a function of position and time, respectively. By introducing a loss function to quantify the deviation between theory and experiment, the pulse characterization ultimately aims at minimizing the loss function with respect to the assumed set of pulse parameters. This approach has already proven its general capability for characterizing the spatial fluence profile [21], but up to now has not been applied to calibrating the pulse duration. To this end, we take advantage of CSD measurements at several different pulse energies, in contrast with the conventional calibration conducted at a fixed pulse energy.

Numerical optimization of black-box functions, i.e., functions that only provide the functional value itself but lack further information such as the gradient, can be accomplished by a huge variety of different algorithms, referred to as meta-heuristic methods [58]. Prominent examples are simulated annealing [59,60] and derivatives thereof [61–64],

particle swarm optimization [65–67], and differential evolution [68,69]. To obtain an appropriate approximation of the function's extremum, however, the listed methods generally require a large number of function evaluations on the considered domain, usually drawn at random from a predefined probability distribution, such that, statistically, the whole domain is sufficiently explored. In spite of the efficiency of XATOM, the computational cost of simulating the ionization dynamics with a huge amount of different parameter sets, hence evaluating the loss function in a high-dimensional parameter space, renders the application of meta-heuristic optimization techniques unfeasible.

In contrast, our proposed XFEL calibration builds on Bayesian optimization (BO) [70–72]. Based on a continuously updating machine-learning model, BO makes a sophisticated prediction where it is most beneficial to search for the function's optimum, rather than drawing random samples from the domain. The obtained function value at the proposed position, in turn, eventually improves the underlying model and the accompanying quality of the predicted extremum. By incorporating the obtained knowledge on the objective function in the numerical optimization, BO has proven its capability to accurately and efficiently solve optimization problems, particularly when a rigorous exploration of the search domain is hampered by the computational costs of the objective function. As such, it is applicable to a huge variety of scientific fields, including hyperparameter tuning of machine-learning and deep-learning models [73,74], chemical synthesis guidance [75], prediction and customization of material properties [76–78], solving inverse problems in physics and model prediction [79,80], preparation and control of ultracold atoms [81–83], and optimization and control of accelerators [84] and free-electron lasers [85] (see Ref. [70] and references therein for a more extensive review on BO applications).

Throughout this work, we perform the XFEL pulse calibration based on experimental data recorded at the Small Quantum Systems (SQS) scientific instrument at the European XFEL [54,86–89] in connection with the experiment by LaForge *et al.* [90]. Our results show that the proposed calibration scheme of utilizing experimental CSDs of light noble gas atoms at a series of pulse energies along with their theoretical counterpart is well suited for determining both the pulse duration and spatial fluence distribution of intense XFEL pulses subsequent to the actual experiment. The accompanying optimization task is efficiently and accurately solved using Bayesian optimization.

The structure of this manuscript is as follows: In Sec. II, we introduce the loss function required for calibration and provide the needed theoretical background. Bayesian optimization as the backbone of our calibration is described in Sec. III. We provide an in-depth discussion of the obtained results in Sec. IV, before concluding this work in Sec. V.

## II. CALIBRATION OUTLINE

The proposed scheme for the calibration of intense XFEL pulses builds on a quantitative comparison of experimental and theoretical CSDs of light noble gas atoms, e.g., Ne. In contrast to previous work on XFEL calibration [21], where the

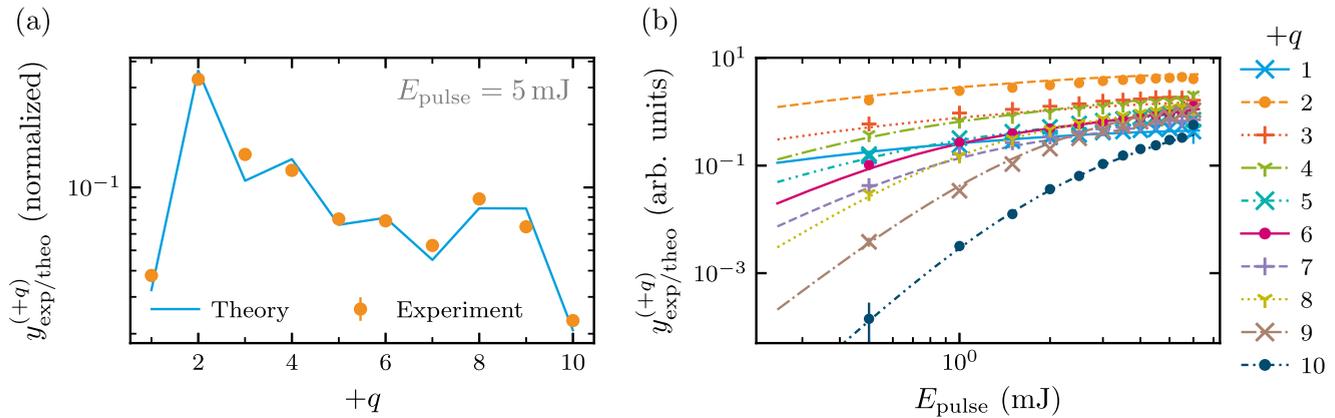


FIG. 1. Experimental and theoretical CSDs (ion yield as a function of charge state) of Ne at a photon energy of 1550 eV with two calibration schemes: (a) calibration at a fixed pulse energy and (b) calibration using a series of pulse energies (present work). The pulse energies in (a) and (b) refer to the nominal experimental pulse energies  $E_i$ . The transmission of the beamline is taken into account in terms of the scaling factor  $s_E$  for computing the theoretical CSDs.

CSD at only one pulse energy was considered [see Fig. 1(a)], we utilize a range of pulse energies and corresponding CSDs, as illustrated in Fig. 1(b). This is enabled by shot-to-shot fluctuations of the pulse energy during the pulse train, which provide a spectrum of CSDs at different pulse-energy bins. We assume that, if no external beam attenuator is employed, the average XFEL pulse shape remains unchanged. Similarly, the use of gas attenuators for varying the pulse energy, which ideally attenuate the pulse energy uniformly, would also provide a CSD spectrum without affecting the pulse shape [86]. The use of solid-foil attenuators, however, considerably affects the spatial fluence distribution [21]. Also we cannot rule out that a drift of the electron bunch during the long pulse train influences various pulse parameters.

Utilizing a series of pulse energies provides more data points for comparison with the theoretical CSDs. In Fig. 1(a), only ten points are available for calibration using Ne, while there are  $10N$  points in Fig. 1(b), where  $N$  is the number of distinct pulse-energy bins extracted from experiment. Including more data points in the calibration is beneficial for resolving subtle dependences of the CSDs on both fluence and pulse duration. Furthermore, the pulse-energy dependence of the ion yields, whose slope in a log-log-graph [Fig. 1(b)] before saturation [21] provides the number of absorbed photons, is naturally incorporated in the calibration procedure. After saturation, the corresponding ion yield needs to be excluded in the conventional one-pulse-energy calibration scheme [21], which hinders the fluence distribution calibration in the high-fluence regime, where many charge states become saturated. The proposed calibration scheme, which systematically covers a range of pulse energies before and after saturation, can overcome this difficulty.

The experimental data set contains a series of ion yields  $y_{\text{exp}}^{(+q)}(E_i)$  for charges  $+q \in Q_s$ , recorded at  $N$  different pulse energies  $E_i$  ( $i = 1, \dots, N$ ). Here,  $Q_s$  denotes the set of selected charges considered for the calibration. To obtain the theoretical counterpart  $y_{\text{theo}}^{(+q)}$ , volume integration constitutes an essential step, incorporating a range of photon fluences provided by the fluence distribution function of the XFEL

pulse. See Ref. [21] for details on volume integration. To characterize the fluence distribution function, we employ a double Gaussian spatial profile (DGSP), in which the spatial pulse shape is described by two Gaussian profiles, giving rise to a narrow high-fluence peak and a wide low-fluence tail [86], in the two dimensions perpendicular to the beam propagation. The dependence of the fluence distribution on the beam propagation direction is neglected with the assumption of the Rayleigh length being longer than the ion detector slit size [21]. In this case, the fluence distribution function is determined by a set of parameters,  $\mathbf{P} = (f_r, w_r, s_E)$ , where  $f_r$  and  $w_r$ , respectively, describe the fluence and width ratios between the first and second Gaussians. The additional parameter  $s_E$  is a scaling factor for the pulse energy, in connection with the ratio of the transmission  $T$  of the x-ray optics and the focal spot size  $\Delta$  of the first Gaussian:  $T/\Delta^2 = s_E(T'/\Delta'^2)$ , where we choose fixed values of  $T' = 1$  and  $\Delta'^2 = 1 \mu\text{m}^2$  in practice. With these parameters, the peak fluence of the DGSP corresponding to the pulse energy  $E_i$  (with a photon energy of  $\omega$ ) is given by

$$F_i = \frac{4 \ln 2}{\pi \omega} \frac{1 + f_r}{1 + f_r w_r^2} \left( \frac{T'}{\Delta'^2} \right) s_E E_i. \quad (1)$$

In this way, the theoretical counterpart of ion yields for  $E_i$  is represented as  $y_{\text{theo}}^{(+q)}(F_i; \mathbf{P})$ . The ansatz of a DGSP is a convenient parametrization of the spatial fluence profile that has been shown to provide volume-integrated ion yields that are in good agreement with experimental results [52,91–93]. Note that there are generally no restrictions for parametrizing the shape of the spatial fluence distribution profile. For instance, it is possible to use the spatial shape from a ray-tracing simulation of the beam or a wavefront sensor measurement [86] with suitable parametrization.

We calibrate XFEL pulses by determining suitable parameters  $\mathbf{P}^* = (f_r^*, w_r^*, s_E^*)$  such that the difference between experiment and theory is minimized. To that end, we introduce the loss function  $L$  as the cumulated logarithmic difference between the experimental ion yields  $y_{\text{exp}}^{(+q)}(E_i)$  and the theoretical

predictions  $y_{\text{theo}}^{(+q)}(F_i; \mathbf{P})$ ,

$$L(\mathbf{P}) = \sum_{i=1}^N \sum_{q \in Q_s} \left[ \log_{10} \frac{y_{\text{theo}}^{(+q)}(F_i; \mathbf{P}) \cdot Y}{y_{\text{exp}}^{(+q)}(E_i)} \right]^2 w_i^{(+q)}. \quad (2)$$

The sums run over all pulse energies and selected charges  $Q_s$ . In principle, the weighting factors  $w_i^{(+q)}$  can be based on the experimental uncertainties,  $\varepsilon_{\text{exp}}^{(+q)}(E_i)$ , for the yields of the measured charge states. For example, one may employ the Gaussian form

$$w_i^{(+q)} = \exp \left[ -\alpha \left( \frac{\varepsilon_{\text{exp}}^{(+q)}(E_i)}{y_{\text{exp}}^{(+q)}(E_i)} \right)^2 \right], \quad (3)$$

where  $\alpha$  is a real coefficient. For the experimental data considered in the present work, we have numerically verified that minimization of  $L$  for  $\alpha = 10$  and 0 (i.e.,  $w_i^{(+q)} = 1$ ) does not give any significant difference (not shown here). Thus, we do not account for the experimental uncertainties in the present work. By log-scaling the ion yields, we ensure that all ion yields contribute equally to the loss function despite spanning several orders of magnitude. To match the different absolute scales of the ion yields, we introduce a global factor  $Y$  that is dynamically determined such that it minimizes  $L$  according to  $\partial L / \partial Y = 0$ ,

$$\log_{10} Y = - \frac{\sum_{i=1}^N \sum_{q \in Q_s} \left[ \log_{10} \frac{y_{\text{theo}}^{(+q)}(F_i; \mathbf{P})}{y_{\text{exp}}^{(+q)}(E_i)} \right] w_i^{(+q)}}{\sum_{i=1}^N \sum_{q \in Q_s} w_i^{(+q)}}. \quad (4)$$

Note that in the previous calibration procedure using one pulse energy [21], as shown in Fig. 1(a), both experimental and theoretical CSDs are individually normalized such that the sum over ion yields equals unity [21], while  $Y$  represents a global normalization factor in the present procedure.

Considering the loss function in Eq. (2), the calibration of intense XFEL pulses boils down to a numerical optimization problem. Here, Eq. (2) represents a black-box function as its evaluation does not provide any additional information on functional properties and, therefore, precludes the use of, e.g., gradient-based optimization methods. Meta-heuristic methods [58] provide a wide variety of algorithms for numerical optimization of such black-box functions. These methods, however, are known to require a large number of function evaluations in order to provide reliable results. This makes their use disadvantageous for the outlined calibration scheme as Eq. (2) involves computationally expensive simulations of x-ray multiphoton ionization dynamics, limiting the number of evaluations available. To overcome this limitation, we employ Bayesian optimization [70–72], a machine-learning-based method that is particularly well suited for numerical optimization of computationally expensive black-box functions.

Powered by Bayesian optimization, which will be explained in the subsequent section, we introduce an additional fit parameter for calibration, the pulse duration  $\tau$ , referring to the full width at half-maximum (FWHM) of the ensemble-averaged temporal profile. The latter is assumed to be a Gaussian function. The CSDs become sensitive to the pulse duration when the pulse duration competes with the lifetime of core-hole states [20,94] or the x-ray intensity is sufficiently high such that the photoionization rate competes with the decay

rate of core-hole states [95]. Then, the parameter space to be optimized becomes four-dimensional:  $\mathbf{P} = (f_r, w_r, s_E, \tau)$ . Such a high-dimensional parametric space is already unfavorable for a brute-force search. For example, if only 100 cases are sampled for each parameter, then the total number of brute-force searches is  $10^8$ . As will be demonstrated in Sec. IV, this number of iterations is dramatically reduced when using Bayesian optimization. The set of parameters considered in this work is of course not unique but depends on the assumed spatial and temporal pulse profile. The outlined calibration, however, can be applied to any set of parameters necessary to appropriately parametrize the pulse profiles.

### III. BAYESIAN OPTIMIZATION

Bayesian optimization (BO) [70–72] is an iterative method for numerically maximizing a target function  $f : U \subset \mathbb{R}^d \rightarrow \mathbb{R}$ , which is assumed to be continuous throughout this work. The key paradigm of BO is not to draw random samples from a probability distribution at which the target function is evaluated, like it is done in simulated annealing [59,60], for instance, but to propose an  $\mathbf{x} \in U$  with the highest expected benefit. The prediction is based on the knowledge of the target function obtained from previous function evaluations. This approach makes BO particularly useful when evaluating  $f$  is computationally expensive, thus limiting the total number of function evaluations. The BO algorithm generally consists of two main components:

(i) *A surrogate model*: Based upon all available function evaluations, a probabilistic model of the function, dubbed surrogate, is built. The surrogate provides a predicted value along with an uncertainty measure for any unknown input  $\mathbf{x} \in U$ . Here, we choose Gaussian processes, the most common surrogate model.

(ii) *An acquisition function*: The acquisition function analyzes the surrogate model and assigns a numerical value encoding the desirability to each point in the domain of  $f$ . The maximum position of the acquisition function indicates the point with the highest expected benefit, which is therefore the most promising subsequent point for evaluating  $f$ . The result, in turn, improves the underlying surrogate model, eventually leading to an updated acquisition and prediction of the maximum.

#### A. Gaussian process regression

Gaussian process regression (GPR) [70,96] is a supervised machine-learning technique, i.e., the attempt of approximating the target function  $f$  based on a collection of  $n$  labeled examples,  $\mathcal{D} = \{(\mathbf{x}_i, y_i) | i = 1, \dots, n\}$ , called training data. The input  $\mathbf{x}_i \in U$  is referred to as a feature vector, and its individual elements  $x_i^{(j)}$  ( $j = 1, \dots, d$ ) are called features. The labels  $y_i$  are considered continuous throughout this work. We further assume  $y_i$  to be noiseless observations of  $f$ , i.e.,  $y_i = f(\mathbf{x}_i)$ . With that, we introduce the more compact notation of an input matrix  $X$ ,

$$X = (\mathbf{x}_1, \dots, \mathbf{x}_n) = \begin{pmatrix} x_1^{(1)} & \cdots & x_n^{(1)} \\ \vdots & \ddots & \vdots \\ x_1^{(d)} & \cdots & x_n^{(d)} \end{pmatrix}, \quad (5)$$

and a corresponding label vector  $\mathbf{y} = (y_1, \dots, y_n)^\top$ .

The fundamental idea of GPR is to assign probabilities to every possible function instead of *a priori* restricting the regression to a certain class of functions, e.g., linear, quadratic, or trigonometric functions. High probabilities indicate functions that are considered to be more likely, also incorporating prior knowledge or assumptions of functional properties. The *prior* distribution of functions represents the beliefs in the kind of functions expected to be observed before seeing any data. The training data  $\mathcal{D}$  restrict the function distribution such that it only contains those functions consistent with the observations. The combination of the prior distribution and the training data yields the *posterior* distribution over functions.

A Gaussian process  $g_f$  for modeling the target function  $f$ , written as [96]

$$g_f(\mathbf{x}) \sim \mathcal{GP}(m(\mathbf{x}), k(\mathbf{x}, \mathbf{x}')), \quad (6)$$

is a collection of random variables, any finite number of which have a joint Gaussian (normal) distribution. It is entirely specified by a prior mean function  $m : U \rightarrow \mathbb{R}$  and a positive semidefinite covariance (kernel) function  $k : U \times U \rightarrow \mathbb{R}$ ,

$$m(\mathbf{x}) = \mathbb{E}[g_f(\mathbf{x})], \quad (7)$$

$$\begin{aligned} k(\mathbf{x}, \mathbf{x}') &= \text{cov}(g_f(\mathbf{x}), g_f(\mathbf{x}')) \\ &= \mathbb{E}[(g_f(\mathbf{x}) - m(\mathbf{x}))(g_f(\mathbf{x}') - m(\mathbf{x}'))], \end{aligned} \quad (8)$$

with  $\mathbb{E}[\cdot]$  and  $\text{cov}(\cdot, \cdot)$  denoting the expectation value and the covariance of two random variables, respectively. For a given covariance function, Eq. (8) enables an explicit formulation of the cross-covariance matrix,

$$K(X, X') = \begin{pmatrix} k(\mathbf{x}_1, \mathbf{x}'_1) & \cdots & k(\mathbf{x}_1, \mathbf{x}'_m) \\ \vdots & \ddots & \vdots \\ k(\mathbf{x}_n, \mathbf{x}'_1) & \cdots & k(\mathbf{x}_n, \mathbf{x}'_m) \end{pmatrix}, \quad (9)$$

with  $X = (\mathbf{x}_1, \dots, \mathbf{x}_n)$  and  $X' = (\mathbf{x}'_1, \dots, \mathbf{x}'_m)$  denoting two input matrices, as defined in Eq. (5). For  $X = X'$ , Eq. (9) gives the covariance matrix, with the diagonal entries being the variances,  $\text{var}(g_f(\mathbf{x})) = \text{cov}(g_f(\mathbf{x}), g_f(\mathbf{x}))$ . Throughout this work, we follow the common assumption of a zero-mean and unit-variance,  $\text{var}(g_f(\mathbf{x})) = k(\mathbf{x}, \mathbf{x}) = 1$ , of the prior distribution, and we incorporate this assumption by normalizing the labels when determining the posterior distribution. However, this does not imply the function values of the prior distribution or the predicted mean values on unseen data are zero.

Evaluating the Gaussian process for an arbitrary feature  $\mathbf{x}$  before taking into account the labels  $\mathbf{y}$  of the training data yields a random number  $g_f|\mathbf{x}$ , which follows the *prior* normal distribution,

$$g_f|\mathbf{x} \sim \mathcal{N}(0, K(\mathbf{x}, \mathbf{x})). \quad (10)$$

Similarly, when applied to the training features  $X$ , the Gaussian process gives a random vector  $\mathbf{g}_f = (g_f(\mathbf{x}_1), \dots, g_f(\mathbf{x}_n))$ , following a multivariate normal distribution,

$$\mathbf{g}_f|X \sim \mathcal{N}(0, K(X, X)). \quad (11)$$

The joint normal distribution of the priors  $\mathbf{g}_f$  and  $g_f$  defines the joint prior distribution of the Gaussian process [96],

$$\begin{bmatrix} \mathbf{g}_f \\ g_f \end{bmatrix} \sim \mathcal{N}\left(0, \begin{bmatrix} K(X, X) & K(X, \mathbf{x}) \\ K(\mathbf{x}, X) & K(\mathbf{x}, \mathbf{x}) \end{bmatrix}\right). \quad (12)$$

The *posterior* distribution is obtained by restricting the joint prior distribution such that it only contains functions that are in agreement with the observed training data. This corresponds to conditioning Eq. (12) on the observations [96], i.e.,  $\mathbf{g}_f = \mathbf{y}$  and

$$g_f|\mathbf{x}, X, \mathbf{y} \sim \mathcal{N}(\mu(\mathbf{x}), \sigma^2(\mathbf{x})), \quad (13)$$

with  $\mu$  and  $\sigma^2$ , respectively, denoting the predicted mean,

$$\mu(\mathbf{x}) = \mathbb{E}[g_f|\mathbf{x}, X, \mathbf{y}] = K(\mathbf{x}, X)K(X, X)^{-1}\mathbf{y}, \quad (14)$$

and the predicted variance,

$$\begin{aligned} \sigma^2(\mathbf{x}) &= \text{var}(g_f|\mathbf{x}, X, \mathbf{y}) \\ &= K(\mathbf{x}, \mathbf{x}) - K(\mathbf{x}, X)K(X, X)^{-1}K(X, \mathbf{x}). \end{aligned} \quad (15)$$

Equations (14) and (15) are the central quantities of the GPR and an integral part of the BO algorithm: The mean predicts the value of the target function  $f$  at each input point  $\mathbf{x}$ , while the variance provides an uncertainty measure of the prediction. By construction, GPR exactly reproduces the training data, i.e.,  $\mu(\mathbf{x}_i) = y_i$  and  $\sigma^2(\mathbf{x}_i) = 0$  ( $i = 1, \dots, n$ ).

Through Eq. (9), the quality of the predictions using GPR strongly depends on the choice made for the covariance (kernel) function  $k$  and its parameters. The squared exponential (SE) covariance function constitutes the most common choice in the context of BO and is used throughout this work (the SE covariance is sometimes referred to as the radial basis function, although this term applies to isotropic kernels in general) [70,72,96]:

$$k_{\text{SE}}(\mathbf{x}, \mathbf{x}') = \exp\left(-\frac{r(\mathbf{x}, \mathbf{x}')^2}{2}\right). \quad (16)$$

Here,  $r(\mathbf{x}, \mathbf{x}')$  is the anisotropic distance between two feature vectors,  $\mathbf{x}$  and  $\mathbf{x}'$ ,

$$r(\mathbf{x}, \mathbf{x}') = \sqrt{(\mathbf{x} - \mathbf{x}')^\top M(\mathbf{x} - \mathbf{x}')}, \quad (17)$$

$M$  being a  $d \times d$  diagonal matrix containing the squared characteristic inverse length scales  $\theta_i^2$  ( $i = 1, \dots, d$ ). Therefore, the GPR with a SE covariance kernel contains  $d$  model parameters, denoted  $\boldsymbol{\theta} = (\theta_1, \dots, \theta_d)$ , which are chosen such that they maximize the probability of the model parameters based on the given training data  $p(\boldsymbol{\theta}|\mathbf{y}, X)$ . Here, the label vector  $\mathbf{y}$  is treated as a random sample drawn from the multivariate normal distribution in Eq. (11). Using Bayes' theorem and the fact that the choice of  $\boldsymbol{\theta}$  is independent of the training input  $X$ , i.e.,  $p(\boldsymbol{\theta}|X) = p(\boldsymbol{\theta})$ , the probability of the kernel parameters  $\boldsymbol{\theta}$  conditioned on the training data is given by

$$p(\boldsymbol{\theta}|\mathbf{y}, X) = \frac{p(\mathbf{y}|X, \boldsymbol{\theta}) p(\boldsymbol{\theta})}{p(\mathbf{y}|X)}. \quad (18)$$

This equation is approximated by the likelihood  $p(\mathbf{y}|X, \boldsymbol{\theta})$ , describing the probability to obtain the labels  $\mathbf{y}$  by randomly sampling from the probability distribution in Eq. (11) at the input  $X$ , evaluated for a specific choice of  $\boldsymbol{\theta}$ . By definition of

the Gaussian process, the likelihood follows a normal distribution [96]:

$$\ln p(\mathbf{y}|X, \boldsymbol{\theta}) = -\frac{1}{2}\mathbf{y}^\top K(X, X)^{-1}\mathbf{y} - \frac{1}{2}\ln |K(X, X)| - \frac{d}{2}\ln 2\pi. \quad (19)$$

The likelihood  $p(\mathbf{y}|X, \boldsymbol{\theta})$  implicitly depends on  $\boldsymbol{\theta}$  through the definition of the covariance matrix in Eq. (9) and the specific choice of the SE covariance function in Eqs. (16) and (17). Finally, the optimal kernel parameters  $\boldsymbol{\theta}^*$ , indicating the model that is most likely to produce the labels  $\mathbf{y}$  at the input  $X$ , are determined by maximizing the (logarithmic) likelihood,

$$\boldsymbol{\theta}^* = \underset{\boldsymbol{\theta}}{\operatorname{argmax}} \ln p(\mathbf{y}|X, \boldsymbol{\theta}). \quad (20)$$

When updating the training data  $\mathcal{D}$  by adding additional feature vectors and corresponding labels, Eq. (20) has to be solved numerically, yielding an updated best model.

### B. Acquisition function

Beside the surrogate model, the acquisition function  $\alpha$  constitutes the main component of BO as it defines the search strategy for finding the maximum of the target function. A vast number of different approaches for constructing both effective and efficient acquisition functions exist, categorized in different classes such as improvement-based policies, optimistic policies, and information-based policies [70,72]. Here, we choose the common but by no means exclusive upper confidence bound (UCB) acquisition function, categorized as an optimistic policy. Following its guiding principle of being optimistic in the face of uncertainty, it corresponds to effectively using a fixed-probability best-case scenario according to the applied surrogate model [70,72]:

$$\alpha_{\text{UCB}}(\mathbf{x}) = \mu(\mathbf{x}) + \kappa\sigma(\mathbf{x}), \quad (21)$$

where  $\mu$  and  $\sigma$  denote the predicted mean value and corresponding standard deviation of the underlying Gaussian process, given by Eqs. (14) and (15), respectively, and  $\kappa$  constitutes a free hyperparameter. The choice made for  $\kappa \geq 0$  is essential for a well-performing optimization as it incorporates a tradeoff between exploitation and exploration of the search space. Generally, small values favor the exploitation of already known regions in the search domain. This results in quick convergence but comes at the risk of getting stuck in a local maximum. In contrast, large values favor the exploration of lesser known regions increasing the probability of finding the global maximum. However, this comes with slower convergence. An attempt at tuning the optimization with respect to  $\kappa$  is presented in Sec. IV.

As the acquisition is an (optimistic) estimation of the target function, its maximum position indicates the estimated maximum position of the target function and, therefore, it constitutes the most beneficial position for the next function evaluation according to the chosen search strategy. The result,

in turn, is used to update the surrogate model and the acquisition function until a maximum number of function evaluations is reached. This procedure delegates the numerical optimization of the target function  $f$  to two consecutive numerical maximizations in each iteration of the BO, namely training the Gaussian process surrogate model by solving Eq. (20), and suggesting a new feature vector by maximizing Eq. (21) with respect to  $\mathbf{x}$ . The likelihood in Eqs. (19) and (20) as well as the acquisition in Eq. (21) as a composition of the predicted mean and variance in Eqs. (14) and (15), respectively, are fully expressible in terms of the kernel function through the definition of the cross-covariance matrix in Eq. (9). The explicit choice of the SE kernel function in Eqs. (16) and (17) therefore provides algebraic and differentiable expressions for the likelihood and acquisition as functions of the model parameters  $\boldsymbol{\theta}$  and feature vectors  $\mathbf{x}$ , respectively, which enables an efficient numerical optimization of Eqs. (20) and (21). For both tasks, we employ the limited-memory Broyden-Fletcher-Goldfarb-Shanno algorithm with bound constraints (L-BFGS-B), as provided by the SCIPY library [97].

Note that the presented implementation of BO is designed for numerical maximization. The XFEL calibration, however, builds on minimizing the loss function in Eq. (2). Therefore, for calibration of the x-ray pulse, we maximize the negative loss. For the sake of simplicity and clarity, we still refer to minimizing the loss function throughout this paper.

## IV. XFEL PULSE CALIBRATION

In the following, we perform an XFEL-pulse calibration using experimental data obtained by LaForge *et al.* [90] at the small quantum systems (SQS) scientific instrument at the European XFEL [54,86–89]. The experiment was conducted with a photon energy of  $\omega = 1550$  eV. The pulse duration deduced from the width of the electron bunch was 25 fs, although the analysis indicates an actual pulse duration of  $\sim 10$  fs full width at half-maximum (FWHM) [90]. The investigated data set consists of ion yields for Ne probed at 12 different pulse energies between 0.5 and 6 mJ measured by an x-ray gas monitor detector [98]. The pulse energy varied due to shot-to-shot fluctuations within the pulse train, which ideally do not affect the spatial pulse profile. Alternatively, a gas attenuator can be used in experiment to ensure a constant spatial profile at varying pulse energies [99]. We describe the calibration procedure for Ne in detail. The corresponding CSDs are depicted in Fig. 1(b). In principle, the presented calibration procedure is applicable to Ar as well, and both Ne and Ar have been widely used for determining the spatial fluence profile [21]. However, we found that the Ar data that were obtained in the same experimental run [90] are not suitable for calibrating the pulse duration because the CSDs of Ar are not sensitive to variations of the pulse duration at the considered photon energy of 1550 eV (see Appendix A for details).

The theoretical CSDs are obtained by simulating x-ray multiphoton ionization dynamics of Ne interacting with an intense XFEL pulse employing the XATOM [56,57] toolkit. Computational details of the XATOM simulations performed are given in Appendix B.

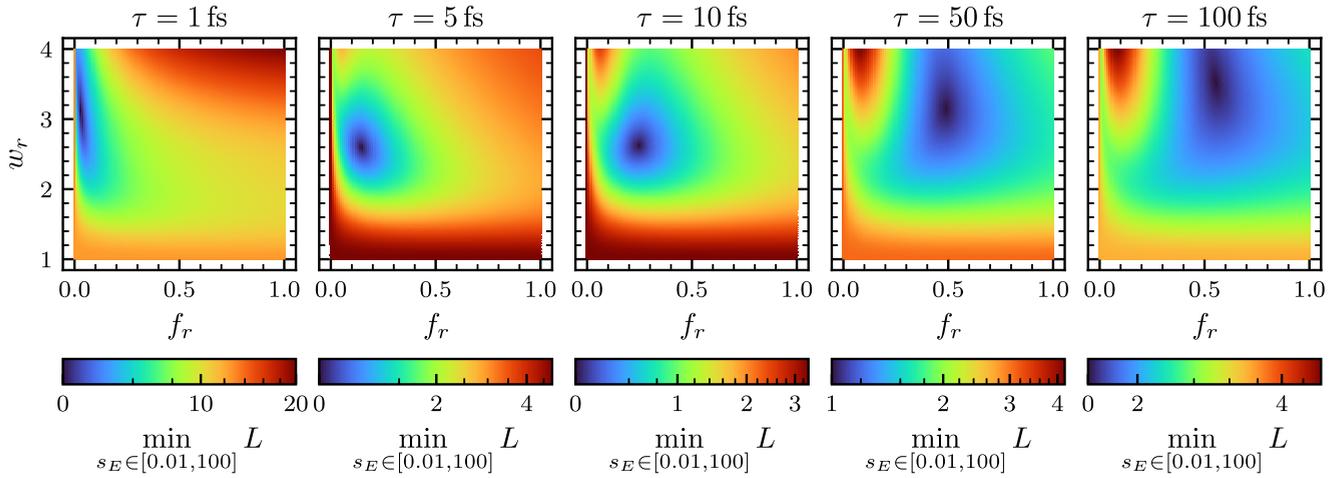


FIG. 2. Manual search for the loss function’s minimum at different but fixed pulse durations  $\tau$ . The pulse parameters  $f_r, w_r$  characterizing the spatial profile are scanned with a respective step size of 0.01.  $s_E$  is chosen such that it minimizes the loss for each point of  $(f_r, w_r)$  and  $\tau$ . The minima and corresponding pulse parameters determined in this way are listed in Table I.

**A. Tuning the BO search algorithm**

Before turning to the calibration of the full four-dimensional parameter space including the pulse duration, we tune the BO algorithm by varying  $\kappa$ , and further test the accuracy and efficiency of our approach to determine the spatial profile of an XFEL pulse from experimental CSD data. To that end, we perform a brute-force minimization of the loss function for several different but fixed pulse durations  $\tau$  by manually scanning the parameter space, and we compare the results with those obtained with BO. We define the search space for both the manual search and the BO as  $S' = [0, 1] \times [1, 4] \times [0.01, 100]$  such that  $(f_r, w_r, s_E) \in S'$ . To obtain a good approximation of the loss function’s minimum, we vary  $(f_r, w_r)$  on a  $(101 \times 301)$ -sized grid. To make it computationally feasible, we do not manually vary  $s_E$  for each  $(f_r, w_r)$ -pair. Instead, we interpolate the respective functional dependence on  $s_E$  using cubic splines based on 50 precalculated data points. Then, we minimize the interpolated loss function with respect to  $s_E$ . The results of this manual search for  $\tau/\text{fs} \in \{1, 5, 10, 50, 100\}$  are depicted in Fig. 2, and the determined minima and calibrated pulse parameters are summarized in Table I. Figure 2 shows that at each pulse duration assumed, the loss function has a minimum, and the optimized parameter set varies for different pulse durations. Such a brute-force search in three-dimensional parameter space is

TABLE I. Manually calibrated spatial pulse parameters and corresponding minimum losses for different fixed pulse durations.

$\tau$ (fs)	$f_r$	$w_r$	$s_E$	$\min_{P' \in S'} L$
1	0.03	3.13	1.18	4.611
5	0.15	2.60	0.37	1.064
10	0.25	2.63	0.33	0.449
50	0.49	3.15	0.42	1.356
100	0.56	3.50	0.51	1.861

computationally expensive and performed only for a few  $\tau$  points to generate reference values.

The UCB acquisition function employed in our BO implementation depends on the hyperparameter  $\kappa$  in Eq. (21), controlling the general search strategy from a more exploring behavior for high values of  $\kappa$  to a dominantly exploiting behavior for small  $\kappa$ . For determining appropriate values for  $\kappa$ , we first perform BO to minimize the loss function  $L(f_r, w_r, s_E, \tau = 10 \text{ fs})$  on the search space  $S'$  for various different  $\kappa \in [0, 4]$ . For each  $\kappa$ , the optimization is repeated 100 times to get statistically significant properties. Each optimization, in turn, starts with two randomly chosen initial points within  $S'$  and terminates with a total of 500 evaluations of the loss function. For comparability between the different settings, we use an identical set of 100 different pairs of initial points for initializing the BO.

To evaluate whether a BO run was successful, we make the following considerations: Given a reference minimum position  $\mathbf{x}_{\text{ref}}^* \in S'$  (obtained by manually scanning the search space) and the current minimum  $\mathbf{x}_N^* \in S'$  after  $N$  iterations, we compute the volume of the three-dimensional ball  $V_{\text{ball}}(r) = \frac{4\pi}{3} r^3$ , with  $r = |\mathbf{x}_{\text{ref}}^* - \mathbf{x}_N^*|$ . Taking into account the different interval lengths, we normalize each coordinate to the unit interval  $[0, 1]$ . Thus, the volume  $\tilde{V}_S$  of the normalized search space equals unity. To quantify the quality of a single optimization run, we introduce  $\eta$  as a measure of by how many orders of magnitude the search space has shrunk with respect to the reference minimum:

$$\eta = -\log_{10} \frac{V_{\text{ball}}(r)}{\tilde{V}_S} = -\log_{10} V_{\text{ball}}(r). \tag{22}$$

A positive value of  $\eta$  indicates an effective reduction of the search space, whereas  $\eta = 0$  if the sphere volume equals the search-space volume. An optimization is considered successful if  $\eta$  related to the current sphere volume is greater than or equal to a predefined threshold value  $\tilde{\eta}$ . Since the manual search for the reference minimum does not provide

the exact position of the minimum but only an approximation based on the choice of the grid, we additionally consider an optimization successful if the loss function is smaller than or equal to the reference value, even if the  $\eta$ -criterion is not satisfied. In practice, this case only occurs if the threshold value is chosen such that the implied distance to the reference minimum is much smaller than the grid spacing of the manual search.

Finally, we define the success rate  $\Sigma^{(\tilde{\eta})}$  as the ratio of successful optimization runs given a threshold  $\tilde{\eta}$  and the total number of optimization runs. Here, we choose  $\tilde{\eta} = 5$  to be sufficiently precise, i.e., the volume of the sphere is five orders of magnitude smaller than the search space. Note that this value corresponds to radii already smaller than the grid spacing used in the manual brute-force search. The success rate  $\Sigma^{(\tilde{\eta}=5)}$  for  $\tau = 10$  fs as a function of the control parameter  $\kappa$  after several different numbers of iterations  $N_{\text{iter}}$  is depicted in Fig. 3(a). For  $\kappa \in [1, 2]$ , 250 iterations are already sufficient to achieve a success rate of  $\sim 95\%$ . The success rate only slightly increases when more iterations are considered. For  $\kappa > 2$ , a similar success rate can be achieved but more iterations are needed. In general,  $\kappa < 1$  gives worse results. Although a high success rate is of course desirable, we note that even with much smaller success rates, it is still likely to find the minimum of the function (or a sufficiently good approximation thereof) since, in practice, the optimization is repeated several times. For instance, if there are 10 optimization runs, a success rate slightly higher than 60% will consequently lead to at least one successful run with a probability of 99.99%. Nonetheless, we recommend choosing  $\kappa \in [1, 2]$  for achieving best results. For the rest of this work, we set  $\kappa = 1.5$ . This choice is validated by the achieved success rates  $\Sigma^{(5)}$  for the remaining four pulse durations, as shown in Fig. 3(b). Like before, the BO starts with two random initial points and is repeated 100 times. Apart from minor deviations during the phase when the success rate rapidly increases, this choice of  $\kappa$  performs equally well in terms of the achieved success rate.

### B. Full pulse calibration

We now turn to the full pulse calibration, including the pulse duration  $\tau$  as an additional pulse parameter to be optimized. The search space  $S$  is four-dimensional such that  $(f_r, w_r, s_E, \tau) \in S = S' \times [1 \text{ fs}, 100 \text{ fs}]$ . The calibration is repeated 25 times, with two random initial points from the search domain initializing each run. The maximum number of iterations is increased to 1500. Based on the results of the 3D calibration in the previous subsection, we set  $\kappa = 1.5$ .

The result of the full four-dimensional calibration, i.e., the calibration giving the smallest loss function, yields the calibrated pulse parameters

$$\mathbf{P}^* = (f_r^*, w_r^*, s_E^*, \tau^*) = (0.262, 2.64, 0.32, 10.9 \text{ fs}).$$

The corresponding theoretical ion yields as a function of the charge in comparison with the experimental CSDs are shown in Fig. 4, which depicts the same data as shown in Fig. 1(b), but using a different representation. The overall agreement between experiment and theory is, apart from small deviations, very good. In particular, the calibrated pulse duration of  $\sim 10$  fs is comparable to the estimated value from the

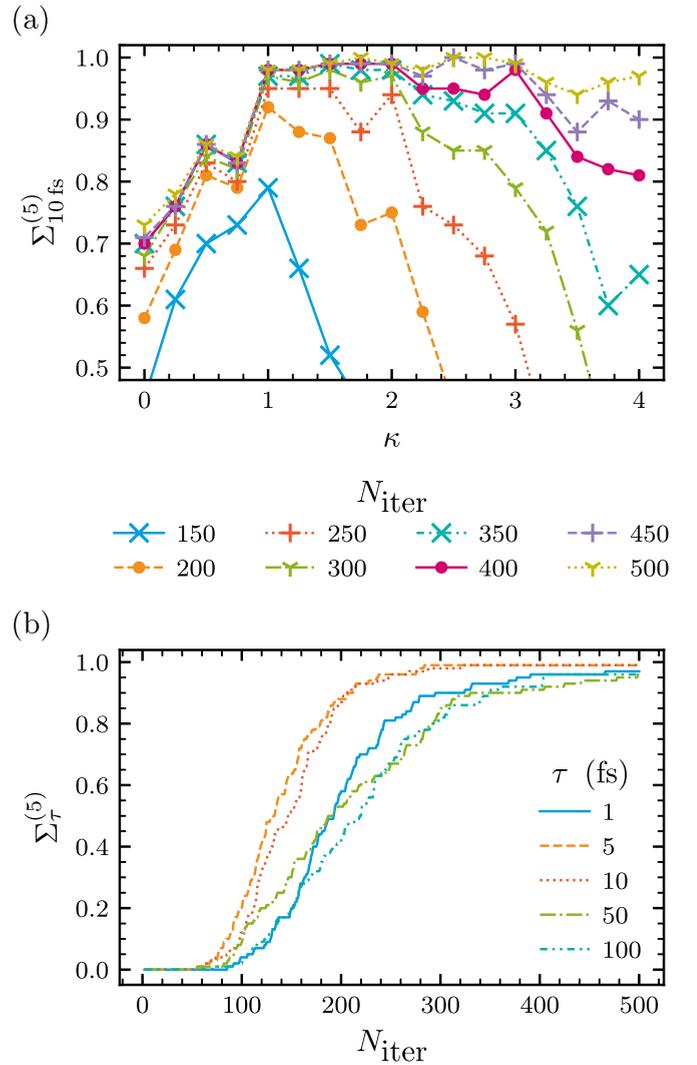


FIG. 3. Tuning of the control parameter  $\kappa$  of the BO algorithm for the three-dimensional pulse calibration with fixed pulse duration  $\tau$ . (a) Success rate of finding the manually determined loss function minimum for  $\tau = 10$  fs as a function of  $\kappa$  after different numbers of iterations  $N_{\text{iter}}$ . Here, a successful optimization run is characterized by a three-dimensional sphere with  $\tilde{\eta} = 5$  (see the main text). (b) Success rate for different pulse durations as a function of the number of iterations. Here, the control parameter is set to  $\kappa = 1.5$ .

x-ray resonance-enhanced multiphoton ionization experiment conducted at the European XFEL [90].

Since a calibration of the spatial pulse profile based on experimental CSDs has already been realized [21,86], our main emphasis is the additional determination of the pulse duration. To verify our procedure, we examine the pulse-duration dependence of the loss function's minimum by manually scanning the search interval [1 fs, 100 fs] and calibrating the spatial pulse profile for each pulse duration, in a similar fashion as for the fine-tuning of  $\kappa$ . For this semimanual three-dimensional calibration of  $(f_r, w_r, s_E)$ , we keep  $\kappa = 1.5$  and a maximum number of iterations of 300. Each three-dimensional spatial profile calibration is repeated ten times. The obtained minimal loss as a function of  $\tau$  is depicted

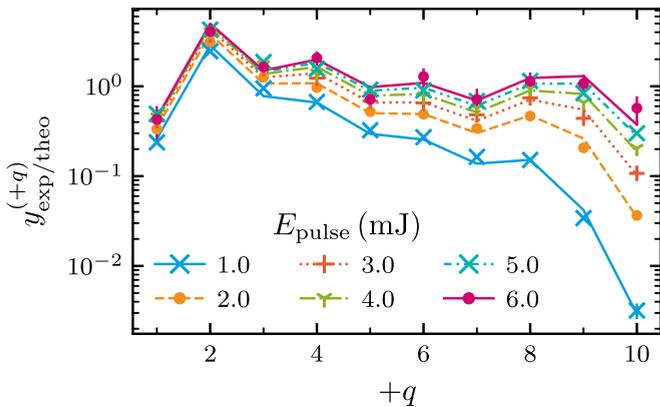


FIG. 4. Experimental and theoretical CSDs, using the calibrated pulse parameters  $(f_r^*, w_r^*, s_E^*, \tau^*) = (0.262, 2.64, 0.32, 10.9 \text{ fs})$ , for various pulse energies. Markers indicate experimental data; lines represent theoretical results.

in Fig. 5. The pulse duration yielding a minimal loss is in perfect agreement with the result obtained in the full four-dimensional calibration.

Having verified the calibrated pulse parameters, we now analyze the general convergence behavior taking into account the results of all 25 performed calibrations. The final pulse parameters of every run are shown in Fig. 6(a). Each marker represents the obtained pulse parameters of one optimization, projected onto two 2D subspaces of  $(f_r, w_r)$  and  $(\tau, s_E)$ . Here, the displayed parameter ranges, containing the results of all calibration runs, correspond to a volume that is approximately  $10^8$  times smaller than the originally considered search-space volume. To further quantify this accumulation of the obtained pulse parameters, we investigate the spread of the individual parameters in terms of the normalized standard deviation  $\sigma_{\text{norm}}$ . Here, “normalized” refers to the fact that all coordinates are normalized to the unit interval taking into account

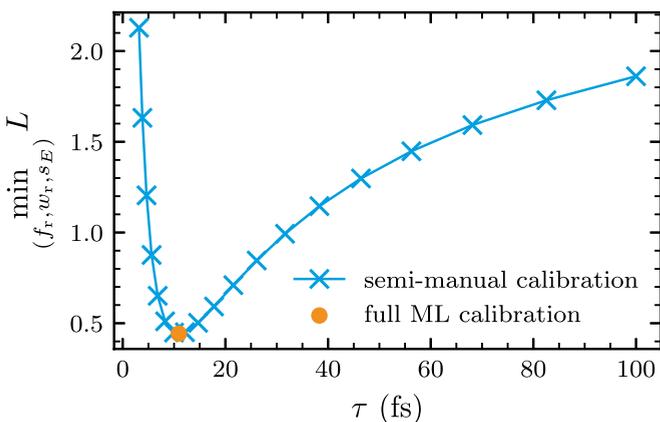


FIG. 5. Minimum losses as a function of pulse duration obtained from the semimanual calibration. For each  $\tau$  considered, the loss function is minimized with respect to the remaining three-dimensional parameter set  $(f_r, w_r, s_E)$  using BO. Also shown is the result of the full four-dimensional calibration.

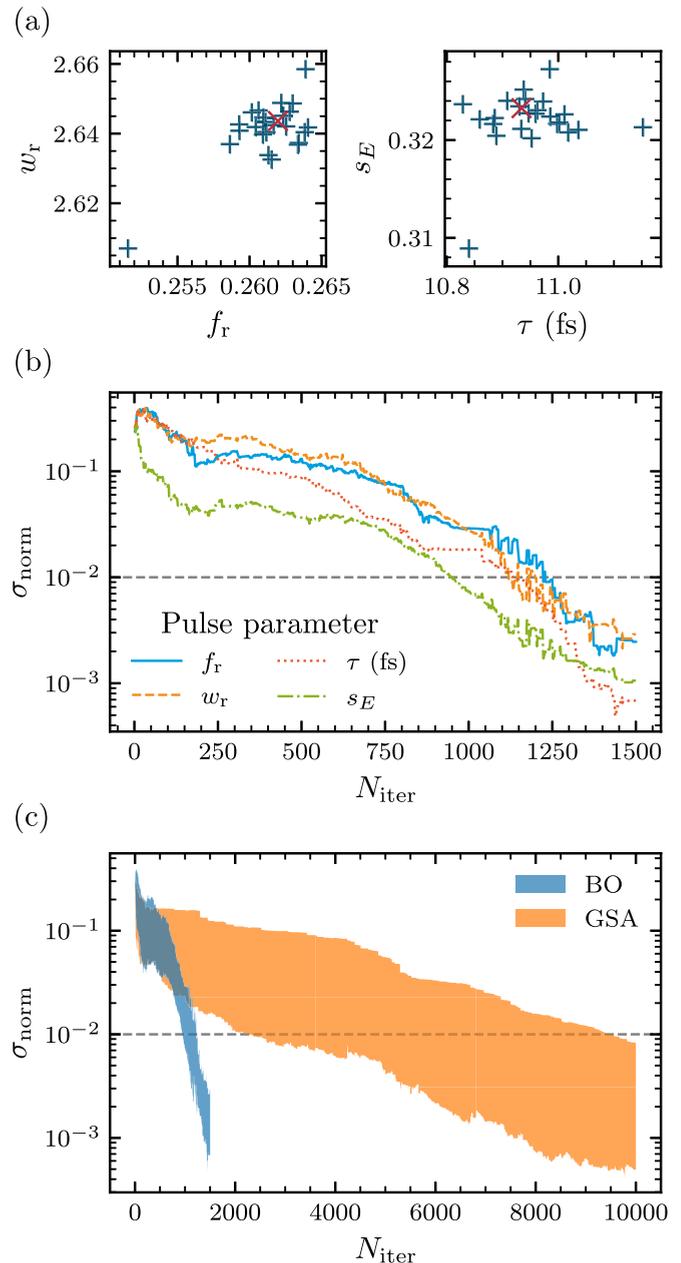


FIG. 6. Results of the four-dimensional pulse calibration, including the pulse duration  $\tau$ . (a) Projection of the calibrated pulse parameters onto the  $(f_r, w_r)$ -plane and the  $(\tau, s_E)$ -plane. Each marker indicates the result of one of the 25 calibration runs; the overall minimum is marked in red. All calibration results lie within a volume that is  $\sim 10^8$  times smaller than the original search-space volume. (b) Normalized standard deviation  $\sigma_{\text{norm}}$  of each determined pulse parameter as a function of the number of iterations considering all 25 BO calibrations. (c) Comparison of the normalized standard deviation  $\sigma_{\text{norm}}$  for BO and generalized simulated annealing (GSA). The colored area marks the spread between the smallest and largest standard deviation out of the four pulse parameters considered. The horizontal dashed lines in (b) and (c) indicate the targeted 1% threshold with respect to the individual search intervals.

the different interval lengths. Naturally, a small  $\sigma_{\text{norm}}$  indicates convergence of the calibration towards a small region in search space. Of course, this does not guarantee that the

true minimum is located inside this region of the search space. It does, however, provide a measure of confidence regarding how likely it is for the calibrated pulse parameters to be located in that specific region. For our purpose, we consider  $\sigma_{\text{norm}} = 1\%$  as a significant and desirable threshold for the calibrated parameters. For each of the four considered pulse parameters, this threshold is passed after  $\sim 1250$  or fewer iterations, as shown in Fig. 6(b).

Finally, we compare the observed convergence behavior of BO to the performance of generalized simulated annealing (GSA) [62,63] as an example of a well-established and frequently used meta-heuristic method. For this comparison, we employ the GSA implementation provided by the SCIPY library [97]. As for BO, the optimization is repeated 25 times started from random points in the search domain (here, only a single point is needed to initialize the optimization). The maximum number of iterations, however, is raised to 10 000. As before, the normalized standard deviation  $\sigma_{\text{norm}}$  is calculated for each pulse parameter individually. The comparison between the BO and GSA results is illustrated in Fig. 6(c). Here, the spread between the smallest  $\sigma_{\text{norm}}$  and the largest value is indicated by colored areas. For GSA, the spread of  $\sigma_{\text{norm}}$  is significantly larger and it takes many more iterations than for BO to reach the desired threshold ( $\sigma_{\text{norm}} = 1\%$ ). Here, we used the default settings of the SCIPY implementation of GSA without fine-tuning.

## V. CONCLUSION

The calibration of intense x-ray free-electron-laser pulses constitutes a challenging task. Yet, the precise characterization of the x-ray pulse is essential for quantitatively comparing experimental and theoretical results. We have proposed a calibration scheme that combines charge-state distributions of light noble gas atoms experimentally determined at a series of pulse energies and their simulated equivalents based on *ab initio* ionization dynamics calculations. Using Bayesian optimization based on Gaussian process regression, we have determined pulse parameters by minimizing the difference between experimental and theoretical outcomes. We have demonstrated in detail how this approach is capable of determining the spatial pulse profile as well as the pulse duration. This includes a sophisticated fine-tuning of the Bayesian optimization providing instructive expertise for future calibration tasks. In the outlined scheme powered by Bayesian optimization, the pulse parameters to be calibrated can be easily adapted to different models of the spatial and temporal pulse profiles, with the possibility to further increase the number of parameters. A comparison of different pulse profile models, however, lies beyond the scope of this work and remains a future task. The detailed theoretical description and numerical simulation of light-matter interaction involving intense x-ray pulses is still an active field of research and is continuously being improved [100–102], potentially to be considered in future calibrations. However, the increasing computational effort on the one hand and the achievable improvements for the calibra-

tion on the other hand are aspects that have to be considered. Experimental uncertainties that ultimately imply uncertainties of the considered loss function have not been considered in the present work. Besides modifying the loss function through suitable weighting factors, experimental uncertainties can be naturally included in the Bayesian optimization, particularly in the underlying Gaussian process regression, by the assumption of Gaussian noise for these uncertainties. Although the inclusion of experimental uncertainties in the case of the present data set does not significantly affect the obtained parameters of the calibrated XFEL pulse, this remains an open question for future investigations.

Data recorded for the experiment at the European XFEL are available on request [103].

## ACKNOWLEDGMENTS

N. Breckwoldt, M.M., and R.S. acknowledge support from the Cluster of Excellence “CUI: Advanced Imaging of Matter” of the Deutsche Forschungsgemeinschaft (DFG)–EXC 2056–project ID 390715994. A.R. and M.M. acknowledge funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)–SFB-925–Project No. 170620586. N. Berrah, D.M., and A.C.L. (UConn) gratefully acknowledge financial support from the Chemical Sciences, Geosciences and Biosciences Division, Office of Basic Energy Sciences, Office of Science, U.S. Department of Energy, Grant No. DE-SC0012376. We acknowledge European XFEL in Schenefeld, Germany, for provision of x-ray free-electron laser beam time at the SQS instrument, and we would like to thank the staff for their assistance.

## APPENDIX A: CALIBRATION USING AR

In principle, the proposed calibration procedure is applicable to CSDs of other light noble gases. Here, we tested the calibration with Ar data obtained from the same experiment as the Ne data [90]. However, the obtained pulse duration of  $\sim 80$  fs differs significantly from the calibrated pulse duration of 10.9 fs obtained from the calibration using Ne. To identify the origin of this discrepancy, we performed a series of three-dimensional calibrations of  $(f_r, w_r, s_E)$  at different pulse durations, analogously to the procedure for Fig. 5. Figure 7(a) shows the minimum losses as a function of pulse duration for Ar, confirming the calibrated pulse duration of about 80 fs. However, note that the range of minimum losses in Fig. 7(a) is much smaller than the range in Fig. 5, making the calibration less reliable. This low sensitivity of the loss function to variations in the pulse duration stems from the fact that the Ar CSDs at the considered photon energy of 1550 eV are largely immune to the pulse duration, as depicted exemplarily for  $E_{\text{pulse}} = 6$  mJ in Fig. 7(b). In contrast, the Ne CSDs exhibit a higher sensitivity to the pulse duration, particularly well pronounced for the ion yield of  $\text{Ne}^{10+}$ , as

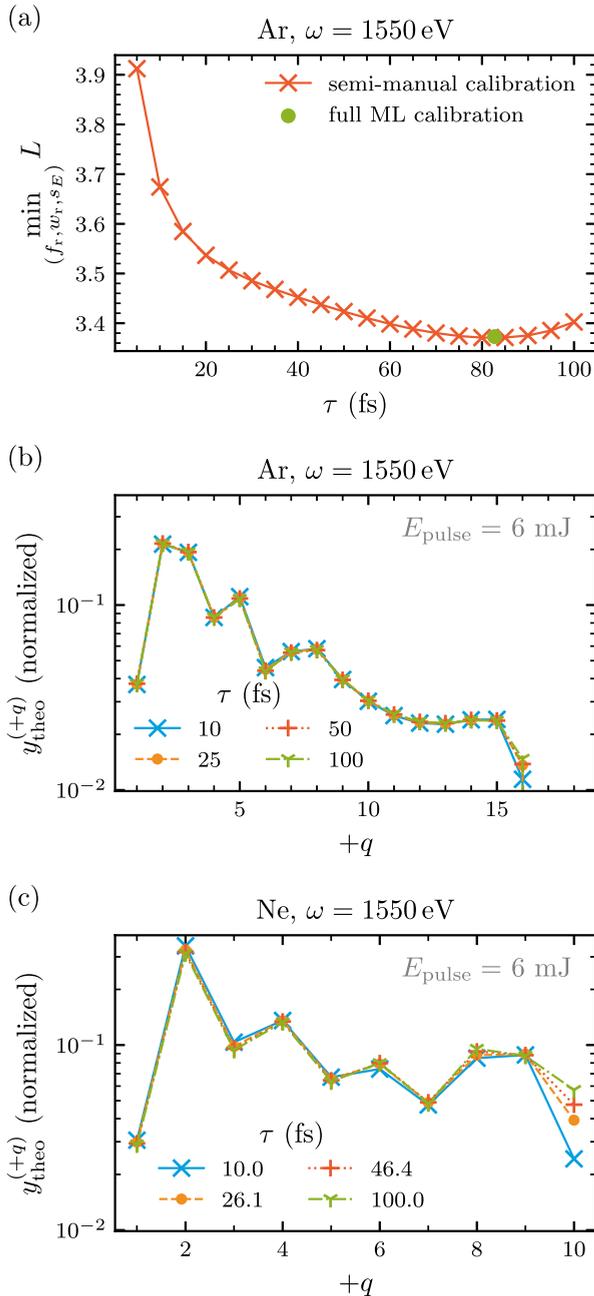


FIG. 7. (a) Minimum losses as a function of pulse duration obtained from the semimanual calibration using Ar data. Analogously to the procedure for Ne described in the main text, the loss function is minimized with respect to the three-dimensional parameter set  $(f_r, w_r, s_E)$  for each pulse duration considered. Also shown is the result of the full four-dimensional calibration. (b) Simulated CSD of Ar at a photon energy of 1550 eV and a pulse energy of 6 mJ for different pulse durations, using the respective calibrated pulse parameters  $(f_r^*, w_r^*, s_E^*)$ . (c) Same as (b) but for Ne.

shown in Fig. 7(c). Additionally, the ion yield of  $\text{Ar}^{16+}$  has to be excluded from the calibration, because it is sensitive to second- or higher-harmonic components in the photon-energy spectrum of the XFEL beam [90]. The remaining ion yields of Ar, including charges from  $q = +1$  to  $+15$ , are virtually independent of the pulse duration, resulting in a rather flat distribution, as shown in Fig. 7(a). Therefore, we conclude that Ar CSDs, at least at a photon energy of 1550 eV, are not suitable for calibrating the pulse duration. We further suggest to examine the dependencies of the theoretical CSDs on all pulse parameters considered before performing a calibration using experimental data.

## APPENDIX B: XATOM CALCULATIONS

All simulated CSDs used for the reported calibrations were obtained using the XATOM toolkit. In XATOM [56,57], the Schrödinger equation is solved within the Hartree-Fock-Slater model. By imposing spherical symmetry on the electron density, each Hartree-Fock-Slater one-electron wave function may be written as a product of a radial function and a spherical harmonic. For bound states, the radial wave functions for each orbital-angular-momentum quantum number  $l$  considered are computed by constructing a matrix representation of the Hartree-Fock-Slater mean-field Hamiltonian using the generalized pseudospectral method [104,105] on a nonuniform grid and then numerically diagonalizing the resulting real-symmetric matrix. Relevant computational parameters are the number of grid points  $N$ , a mapping parameter  $L$ , and the maximum radius  $r_{\text{max}}$ . For continuum states, the radial functions are computed by numerically integrating, at each one-electron energy considered, the radial mean-field Schrödinger equation using the fourth-order Runge-Kutta method on a uniform grid [106,107] with a radial step size  $dr$ . In our calculations,  $N = 200$ ,  $L = 1$ ,  $r_{\text{max}} = 50$  a.u., and  $dr = 0.005$  a.u. are used.

To simulate ionization dynamics, coupled rate equations [20,108,109] are solved using the fourth-order Runge-Kutta method. In our calculations, we treat x-ray-induced photoionization, Auger-Meitner decay, and fluorescence. The associated cross sections and rates are computed using the results of the radial-wave-function calculations described above. In addition, for each process, shakeoff branching ratios are calculated within the sudden approximation [110,111]. The x-ray pulse is assumed to have a Gaussian temporal profile with a duration  $\tau$  (FWHM). The electronic population dynamics triggered by the x-ray pulse are computed for a time interval that is eight times longer than the pulse duration  $\tau$ . The number of temporal grid points used is 20 000. Thus, for  $\tau = 10$  fs, the temporal step size is 0.004 fs.

We confirmed that our results are convergent with respect to all computational parameters.

[1] E. A. Seddon, J. A. Clarke, D. J. Dunning, C. Masciovecchio, C. J. Milne, F. Parmigiani, D. Rugg, J. C. H. Spence, N. R. Thompson, K. Ueda, S. M. Vinko, J. S. Wark, and W. Wurth, Short-wavelength free-electron laser sources and science: A review, *Rep. Prog. Phys.* **80**, 115901 (2017).

[2] C. Pellegrini, A. Marinelli, and S. Reiche, The physics of x-ray free-electron lasers, *Rev. Mod. Phys.* **88**, 015006 (2016).

[3] J. C. H. Spence, XFELs for structure and dynamics in biology, *IUCrJ* **4**, 322 (2017).

- [4] I. Schlichting and J. Miao, Emerging opportunities in structural biology with X-ray free-electron lasers, *Curr. Opin. Struct. Biol.* **22**, 613 (2012).
- [5] K. J. Gaffney and H. N. Chapman, Imaging atomic structure and dynamics with ultrafast x-ray scattering, *Science* **316**, 1444 (2007).
- [6] L. Young, K. Ueda, M. Gühr, P. H. Bucksbaum, M. Simon, S. Mukamel, N. Rohringer, K. C. Prince, C. Masciovecchio, M. Meyer, A. Rudenko, D. Rolles, C. Bostedt, M. Fuchs, D. A. Reis, R. Santra, H. Kapteyn, M. Murnane, H. Ibrahim, F. Légaré *et al.*, Roadmap of ultrafast x-ray atomic and molecular physics, *J. Phys. B* **51**, 032003 (2018).
- [7] L. Fang, T. Osipov, B. F. Murphy, A. Rudenko, D. Rolles, V. S. Petrovic, C. Bostedt, J. D. Bozek, P. H. Bucksbaum, and N. Berrah, Probing ultrafast electronic and molecular dynamics with free-electron lasers, *J. Phys. B* **47**, 124006 (2014).
- [8] C. Bostedt, J. D. Bozek, P. H. Bucksbaum, R. N. Coffee, J. B. Hastings, Z. Huang, R. W. Lee, S. Schorb, J. N. Corlett, P. Denes, P. Emma, R. W. Falcone, R. W. Schoenlein, G. Doumy, E. P. Kanter, B. Kraessig, S. Southworth, L. Young, L. Fang, M. Hoener *et al.*, Ultra-fast and ultra-intense x-ray sciences: first results from the Linac Coherent Light Source free-electron laser, *J. Phys. B* **46**, 164003 (2013).
- [9] J. Ullrich, A. Rudenko, and R. Moshhammer, Free-electron lasers: New avenues in molecular physics and photochemistry, *Annu. Rev. Phys. Chem.* **63**, 635 (2012).
- [10] S. H. Glenzer, L. B. Fletcher, E. Galtier, B. Nagler, R. Alonso-Mori, B. Barbreil, S. B. Brown, D. A. Chapman, Z. Chen, C. B. Curry, F. Fiuza, E. Gamboa, M. Gauthier, D. O. Gericke, A. Gleason, S. Goede, E. Granados, P. Heimann, J. Kim, D. Kraus *et al.*, Matter under extreme conditions experiments at the Linac Coherent Light Source, *J. Phys. B* **49**, 092001 (2016).
- [11] H. N. Chapman, P. Fromme, A. Barty, T. A. White, R. A. Kirian, A. Aquila, M. S. Hunter, J. Schulz, D. P. DePonte, U. Weierstall, R. B. Doak, F. R. N. C. Maia, A. V. Martin, I. Schlichting, L. Lomb, N. Coppola, R. L. Shoeman, S. W. Epp, R. Hartmann, D. Rolles *et al.*, Femtosecond X-ray protein nanocrystallography, *Nature* **470**, 73 (2011).
- [12] E. Sobolev, S. Zolotarev, K. Giewekemeyer, J. Bielecki, K. Okamoto, H. K. N. Reddy, J. Andreasson, K. Ayer, I. Barak, S. Bari, A. Barty, R. Bean, S. Bobkov, H. N. Chapman, G. Chojnowski, B. J. Daurer, K. Dörner, T. Ekeberg, L. Flückiger, O. Galzitskaya *et al.*, Megahertz single-particle imaging at the European XFEL, *Commun. Phys.* **3**, 97 (2020).
- [13] M. M. Seibert, T. Ekeberg, F. R. N. C. Maia, M. Svenda, J. Andreasson, O. Jönsson, D. Odić, B. Iwan, A. Rucker, D. Westphal, M. Hantke, D. P. DePonte, A. Barty, J. Schulz, L. Gumprecht, N. Coppola, A. Aquila, M. Liang, T. A. White, A. Martin *et al.*, Single mimivirus particles intercepted and imaged with an X-ray laser, *Nature* **470**, 78 (2011).
- [14] D. Assalauova, Y. Y. Kim, S. Bobkov, R. Khubbutdinov, M. Rose, R. Alvarez, J. Andreasson, E. Balaur, A. Contreras, H. DeMirici, L. Gelisio, J. Hajdu, M. S. Hunter, R. P. Kurta, H. Li, M. McFadden, R. Nazari, P. Schwander, A. Teslyuk, P. Walter *et al.*, An advanced workflow for single-particle imaging with the limited data at an X-ray free-electron laser, *IUCr J* **7**, 1102 (2020).
- [15] G. Brändén, G. Hammarin, R. Harimoorthy, A. Johansson, D. Arnlund, E. Malmerberg, A. Barty, S. Tångefjord, P. Berntsen, D. P. DePonte, C. Seuring, T. A. White, F. Stellato, R. Bean, K. R. Beyerlein, L. M. G. Chavas, H. Fleckenstein, C. Gati, U. Ghoshdastider, L. Gumprecht *et al.*, Coherent diffractive imaging of microtubules using an X-ray laser, *Nat. Commun.* **10**, 2589 (2019).
- [16] A. Hosseinizadeh, G. Mashayekhi, J. Copperman, P. Schwander, A. Dashti, R. Sepehr, R. Fung, M. Schmidt, C. H. Yoon, B. G. Hogue, G. J. Williams, A. Aquila, and A. Ourmazd, Conformational landscape of a virus by single-particle X-ray scattering, *Nat. Methods* **14**, 877 (2017).
- [17] A. Aquila, A. Barty, C. Bostedt, S. Boutet, G. Carini, D. dePonte, P. Drell, S. Doniach, K. H. Downing, T. Earnest, H. Elmlund, V. Elser, M. Gühr, J. Hajdu, J. Hastings, S. P. Hau-Riege, Z. Huang, E. E. Lattman, F. R. N. C. Maia, S. Marchesini *et al.*, The linac coherent light source single particle imaging road map, *Struct. Dyn.* **2**, 041701 (2015).
- [18] B. Ziaja, H. N. Chapman, R. Fäustlin, S. Hau-Riege, Z. Jurek, A. V. Martin, S. Toleikis, F. Wang, E. Weckert, and R. Santra, Limitations of coherent diffractive imaging of single objects due to their damage by intense x-ray radiation, *New J. Phys.* **14**, 115015 (2012).
- [19] A. Barty, R. Soufli, T. McCarville, S. L. Baker, M. J. Pivovarov, P. Stefan, and R. Bionta, Predicting the coherent X-ray wavefront focal properties at the Linac Coherent Light Source (LCLS) X-ray free electron laser, *Opt. Express* **17**, 15508 (2009).
- [20] L. Young, E. P. Kanter, B. Krässig, Y. Li, A. M. March, S. T. Pratt, R. Santra, S. H. Southworth, N. Rohringer, L. F. DiMauro, G. Doumy, C. A. Roedig, N. Berrah, L. Fang, M. Hoener, P. H. Bucksbaum, J. P. Cryan, S. Ghimire, J. M. Glowia, D. A. Reis *et al.*, Femtosecond electronic response of atoms to ultra-intense X-rays, *Nature* **466**, 56 (2010).
- [21] K. Toyota, Z. Jurek, S.-K. Son, H. Fukuzawa, K. Ueda, N. Berrah, B. Rudek, D. Rolles, A. Rudenko, and R. Santra, *xcalib*: A focal spot calibrator for intense X-ray free-electron laser pulses based on the charge state distributions of light atoms, *J. Synch. Radiat.* **26**, 1017 (2019).
- [22] B. Rösner, F. Döring, P. R. Ribič, D. Gauthier, E. Principi, C. Masciovecchio, M. Zangrando, J. Vila-Comamala, G. D. Ninno, and C. David, High resolution beam profiling of X-ray free electron laser radiation by polymer imprint development, *Opt. Express* **25**, 30686 (2017).
- [23] J. Chalupský, P. Boháček, T. Burian, V. Hájková, S. P. Hau-Riege, P. A. Heimann, L. Juha, M. Messerschmidt, S. P. Moeller, B. Nagler, M. Rowen, W. F. Schlotter, M. L. Swiggers, J. J. Turner, and J. Krzywinski, Imprinting a Focused X-Ray Laser Beam to Measure Its Full Spatial Characteristics, *Phys. Rev. Appl.* **4**, 014004 (2015).
- [24] J. Chalupský, P. Bohacek, V. Hajkova, S. P. Hau-Riege, P. A. Heimann, L. Juha, J. Krzywinski, M. Messerschmidt, S. P. Moeller, B. Nagler, M. Rowen, W. F. Schlotter, M. L. Swiggers, and J. J. Turner, Comparing different approaches to characterization of focused X-ray laser beams, *Nucl. Instrum. Methods Phys. Res. Sect. A* **631**, 130 (2011).
- [25] J. Chalupský, J. Krzywinski, L. Juha, V. Hájková, J. Cihelka, T. Burian, L. Vyšín, J. Gaudin, A. Gleason, M. Jurek, A. R. Khorsand, D. Klinger, H. Wabnitz, R. Sobierajski, M. Störmer, K. Tiedtke, and S. Toleikis, Spot size characterization of focused non-Gaussian X-ray laser beams, *Opt. Express* **18**, 27836 (2010).

- [26] J. Chalupský, L. Juha, J. Kuba, J. Cihelka, V. Hájková, S. Koptyaev, J. Krása, A. Velyhan, M. Bergh, C. Coleman, J. Hajdu, R. M. Bionta, H. Chapman, S. P. Hau-Riege, R. A. London, M. Jurek, J. Krzywinski, R. Nietubyc, J. B. Pelka, R. Sobierajski *et al.*, Characteristics of focused soft X-ray free-electron laser beam determined by ablation of organic molecular solids, *Opt. Express* **15**, 6036 (2007).
- [27] B. Keitel, E. Plönjes, S. Kreis, M. Kuhlmann, K. Tiedtke, T. Mey, B. Schäfer, and K. Mann, Hartmann wavefront sensors and their application at FLASH, *J. Synch. Radiat.* **23**, 43 (2016).
- [28] B. Flöter, P. Juranić, S. Kapitzki, B. Keitel, K. Mann, E. Plönjes, B. Schäfer, and K. Tiedtke, EUV Hartmann sensor for wavefront measurements at the Free-electron LASer in Hamburg, *New J. Phys.* **12**, 083015 (2010).
- [29] B. Schäfer, M. Lübbecke, and K. Mann, Hartmann-Shack wave front measurements for real time determination of laser beam propagation parameters, *Rev. Sci. Instrum.* **77**, 053103 (2006).
- [30] S. Le Pape, P. Zeitoun, M. Idir, P. Dhez, J. J. Rocca, and M. François, Electromagnetic-Field Distribution Measurements in the Soft X-Ray Range: Full Characterization of a Soft X-Ray Laser Beam, *Phys. Rev. Lett.* **88**, 183901 (2002).
- [31] Y. Kayser, S. Rutishauser, T. Katayama, H. Ohashi, T. Kameshima, U. Flechsig, M. Yabashi, and C. David, Wavefront metrology measurements at SACLA by means of X-ray grating interferometry, *Opt. Express* **22**, 9004 (2014).
- [32] S. Rutishauser, L. Samoylova, J. Krzywinski, O. Bunk, J. Grünert, H. Sinn, M. Cammarata, D. M. Fritz, and C. David, Exploring the wavefront of hard X-ray free-electron laser radiation, *Nat. Commun.* **3**, 947 (2012).
- [33] B. Nagler, A. Aquila, S. Boutet, E. C. Galtier, A. Hashim, M. S. Hunter, M. Liang, A. E. Sakdinawat, C. G. Schroer, A. Schropp, M. H. Seaberg, F. Seiboth, T. van Driel, Z. Xing, Y. Liu, and H. J. Lee, Focal Spot and Wavefront Sensing of an X-Ray Free Electron laser using Ronchi shearing interferometry, *Sci. Rep.* **7**, 13698 (2017).
- [34] D. Nilsson, F. Uhlén, A. Holmberg, H. M. Hertz, A. Schropp, J. Patommel, R. Hoppe, F. Seiboth, V. Meier, C. G. Schroer, E. Galtier, B. Nagler, H. J. Lee, and U. Vogt, Ronchi test for characterization of nanofocusing optics at a hard x-ray free-electron laser, *Opt. Lett.* **37**, 5046 (2012).
- [35] M. Schneider, C. M. Günther, B. Pfau, F. Capotondi, M. Manfreda, M. Zangrando, N. Mahne, L. Raimondi, E. Pedersoli, D. Naumenko, and S. Eisebitt, In situ single-shot diffractive fluence mapping for X-ray free-electron laser pulses, *Nat. Commun.* **9**, 214 (2018).
- [36] M. Schneider, C. M. Günther, C. von Korff Schmising, B. Pfau, and S. Eisebitt, Curved gratings as an integrated photon fluence monitor in x-ray transmission scattering experiments, *Opt. Express* **24**, 13091 (2016).
- [37] N. D. Loh, D. Starodub, L. Lomb, C. Y. Hampton, A. V. Martin, R. G. Sierra, A. Barty, A. Aquila, J. Schulz, J. Steinbrener, R. L. Shoeman, S. Kassemeyer, C. Bostedt, J. Bozek, S. W. Epp, B. Erk, R. Hartmann, D. Rolles, A. Rudenko, B. Rudek *et al.*, Sensing the wavefront of x-ray free-electron lasers using aerosol spheres, *Opt. Express* **21**, 12385 (2013).
- [38] A. Schropp, R. Hoppe, V. Meier, J. Patommel, F. Seiboth, H. J. Lee, B. Nagler, E. C. Galtier, B. Arnold, U. Zastra, J. B. Hastings, D. Nilsson, F. Uhlén, U. Vogt, H. M. Hertz, and C. G. Schroer, Full spatial characterization of a nanofocused x-ray free-electron laser beam by ptychographic imaging, *Sci. Rep.* **3**, 1633 (2013).
- [39] A. Kobayashi, Y. Sekiguchi, T. Oroguchi, M. Yamamoto, and M. Nakasako, Shot-by-shot characterization of focused X-ray free electron laser pulses, *Sci. Rep.* **8**, 831 (2018).
- [40] A. Singer, U. Lorenz, F. Sorgenfrei, N. Gerasimova, J. Gulden, O. M. Yefanov, R. P. Kurta, A. Shabalin, R. Dronyak, R. Treusch, V. Kocharyan, E. Weckert, W. Wurth, and I. A. Vartanyants, Hanbury Brown–Twiss Interferometry at a Free-Electron Laser, *Phys. Rev. Lett.* **111**, 034802 (2013).
- [41] O. Y. Gorobtsov, G. Mercurio, G. Brenner, U. Lorenz, N. Gerasimova, R. P. Kurta, F. Hieke, P. Skopintsev, I. Zaluzhnyy, S. Lazarev, D. Dzhigaev, M. Rose, A. Singer, W. Wurth, and I. A. Vartanyants, Statistical properties of a free-electron laser revealed by Hanbury Brown–Twiss interferometry, *Phys. Rev. A* **95**, 023843 (2017).
- [42] O. Y. Gorobtsov, N. Mukharamova, S. Lazarev, M. Chollet, D. Zhu, Y. Feng, R. P. Kurta, J.-M. Meijer, G. Williams, M. Sikorski, S. Song, D. Dzhigaev, S. Serkez, A. Singer, A. V. Petukhov, and I. A. Vartanyants, Diffraction based Hanbury Brown and Twiss interferometry at a hard x-ray free-electron laser, *Sci. Rep.* **8**, 2219 (2018).
- [43] O. Y. Gorobtsov, G. Mercurio, F. Capotondi, P. Skopintsev, S. Lazarev, I. A. Zaluzhnyy, M. B. Danailov, M. Dell’Angela, M. Manfreda, E. Pedersoli, L. Giannessi, M. Kiskinova, K. C. Prince, W. Wurth, and I. A. Vartanyants, Seeded X-ray free-electron laser generating radiation with laser statistical properties, *Nat. Commun.* **9**, 4498 (2018).
- [44] I. Inoue, T. Hara, Y. Inubushi, K. Tono, T. Inagaki, T. Katayama, Y. Amemiya, H. Tanaka, and M. Yabashi, X-ray Hanbury Brown–Twiss interferometry for determination of ultrashort electron-bunch duration, *Phys. Rev. Accel. Beams* **21**, 080704 (2018).
- [45] I. A. Vartanyants and R. Khubbutdinov, Theoretical analysis of Hanbury Brown and Twiss interferometry at soft-x-ray free-electron lasers, *Phys. Rev. A* **104**, 023508 (2021).
- [46] R. Khubbutdinov, N. Gerasimova, G. Mercurio, D. As-salauova, J. Carnis, L. Gelisio, L. le Guyader, A. Ignatenko, Y. Y. Kim, B. van Kuiken, R. P. Kurta, D. Lapkin, M. Teichmann, A. Yaroslavtsev, O. Gorobtsov, A. P. Menushenkov, M. Scholz, A. Scherz, and I. A. Vartanyants, High spatial coherence and short pulse duration revealed by the Hanbury Brown and Twiss interferometry at the European XFEL, *Struct. Dyn.* **8**, 044305 (2021).
- [47] Y. Y. Kim, R. Khubbutdinov, J. Carnis, S. Kim, D. Nam, I. Nam, G. Kim, C. H. Shim, H. Yang, M. Cho, C.-K. Min, C. Kim, H.-S. Kang, and I. A. Vartanyants, Statistical analysis of hard X-ray radiation at the PAL-XFEL facility performed by Hanbury Brown and Twiss interferometry, *J. Synch. Radiat.* **29**, 1465 (2022).
- [48] R. Santra and L. Young, Interaction of intense x-ray beams with atoms, in *Synchrotron Light Sources and Free-Electron Lasers: Accelerator Physics, Instrumentation and Science Applications*, edited by E. J. Jaeschke, S. Khan, J. R. Schneider, and J. B. Hastings (Springer International, Cham, 2016), pp. 1233–1260.
- [49] R. Santra, Concepts in x-ray physics, *J. Phys. B* **42**, 023001 (2009).

- [50] B. Rudek, S.-K. Son, L. Foucar, S. W. Epp, B. Erk, R. Hartmann, M. Adolph, R. Andritschke, A. Aquila, N. Berrah, C. Bostedt, J. Bozek, N. Coppola, F. Filsinger, H. Gorke, T. Gorkhover, H. Graafsma, L. Gumprecht, A. Hartmann, G. Hauser *et al.*, Ultra-efficient ionization of heavy atoms by intense X-ray free-electron laser pulses, *Nat. Photon.* **6**, 858 (2012).
- [51] H. Fukuzawa, S.-K. Son, K. Motomura, S. Mondal, K. Nagaya, S. Wada, X.-J. Liu, R. Feifel, T. Tachibana, Y. Ito, M. Kimura, T. Sakai, K. Matsunami, H. Hayashita, J. Kajikawa, P. Johnsson, M. Siano, E. Kukk, B. Rudek, B. Erk *et al.*, Deep Inner-Shell Multiphoton Ionization by Intense X-Ray Free-Electron Laser Pulses, *Phys. Rev. Lett.* **110**, 173005 (2013).
- [52] A. Rudenko, L. Inhester, K. Hanasaki, X. Li, S. J. Robotjazi, B. Erk, R. Boll, K. Toyota, Y. Hao, O. Vendrell, C. Bomme, E. Savelyev, B. Rudek, L. Foucar, S. H. Southworth, C. S. Lehmann, B. Kraessig, T. Marchenko, M. Simon, K. Ueda *et al.*, Femtosecond response of polyatomic molecules to ultra-intense hard X-rays, *Nature* **546**, 129 (2017).
- [53] J. D. Bozek, AMO instrumentation for the LCLS X-ray FEL, *Eur. Phys. J. Spec. Top.* **169**, 129 (2009).
- [54] M. Meyer *et al.* (unpublished).
- [55] V. Lyamayev, Y. Ovcharenko, R. Katzy, M. Devetta, L. Bruder, A. LaForge, M. Mudrich, U. Person, F. Stienkemeier, M. Krikunova, T. Möller, P. Piseri, L. Avaldi, M. Coreno, P. O’Keeffe, P. Bolognesi, M. Alagia, A. Kivimäki, M. D. Fraia, N. B. Brauer *et al.*, A modular end-station for atomic, molecular, and cluster science at the low density matter beamline of FERMI@Elettra, *J. Phys. B* **46**, 164007 (2013).
- [56] S.-K. Son, L. Young, and R. Santra, Impact of hollow-atom formation on coherent x-ray scattering at high intensity, *Phys. Rev. A* **83**, 033402 (2011).
- [57] Z. Jurek, S.-K. Son, B. Ziaja, and R. Santra, *XMDYN* and *XATOM*: Versatile simulation tools for quantitative modeling of X-ray free-electron laser induced dynamics of matter, *J. Appl. Crystallogr.* **49**, 1048 (2016).
- [58] S. Luke, *Essentials of Metaheuristics*, 2nd ed. (Lulu, 2013), available for free at <http://cs.gmu.edu/~sean/book/metaheuristics/>.
- [59] S. Kirkpatrick, C. D. Gelatt Jr., and M. P. Vecchi, Optimization by simulated annealing, *Science* **220**, 671 (1983).
- [60] A. Khachatryan, S. Semenovskaya, and B. Vainshtein, The thermodynamic approach to the structure analysis of crystals, *Acta Cryst. A* **37**, 742 (1981).
- [61] H. Szu and R. Hartley, Fast simulated annealing, *Phys. Lett. A* **122**, 157 (1987).
- [62] C. Tsallis and D. A. Stariolo, Generalized simulated annealing, *Physica A* **233**, 395 (1996).
- [63] I. Andricioaei and J. E. Straub, Generalized simulated annealing algorithms using Tsallis statistics: Application to conformational optimization of a tetrapeptide, *Phys. Rev. E* **53**, R3055 (1996).
- [64] Y. Xiang and X. G. Gong, Efficiency of generalized simulated annealing, *Phys. Rev. E* **62**, 4473 (2000).
- [65] R. Poli, J. Kennedy, and T. Blackwell, Particle swarm optimization, *Swarm Intell.* **1**, 33 (2007).
- [66] Y. Shi and R. Eberhart, A modified particle swarm optimizer, in *Proceedings of the 1998 IEEE International Conference on Evolutionary Computation: IEEE World Congress on Computational Intelligence* (IEEE, Piscataway, NJ, 1998), pp. 69–73.
- [67] J. Kennedy and R. Eberhart, Particle swarm optimization, in *Proceedings of the ICNN’95 - International Conference on Neural Networks* (IEEE, Piscataway, NJ, 1995), Vol. 4, pp. 1942–1948.
- [68] R. Storn and K. Price, Differential evolution—A simple and efficient heuristic for global optimization over continuous spaces, *J. Global Optim.* **11**, 341 (1997).
- [69] R. Storn, On the usage of differential evolution for function optimization, in *Proceedings of the North American Fuzzy Information Processing* (IEEE, Piscataway, NJ, 1996), pp. 519–523.
- [70] R. Garnett, *Bayesian Optimization* (Cambridge University Press, Cambridge, 2022).
- [71] P. I. Frazier, A tutorial on Bayesian optimization, [arXiv:1807.02811](https://arxiv.org/abs/1807.02811).
- [72] B. Shahriari, K. Swersky, Z. Wang, R. P. Adams, and N. de Freitas, Taking the human out of the loop: A review of bayesian optimization, *Proc. IEEE* **104**, 148 (2016).
- [73] R. Turner, D. Eriksson, M. McCourt, J. Kiili, E. Laaksonen, Z. Xu, and I. Guyon, Bayesian optimization is superior to random search for machine learning hyperparameter tuning: Analysis of the black-box optimization challenge 2020, in *Proceedings of the NeurIPS 2020 Competition and Demonstration Track*, edited by H. J. Escalante and K. Hofmann, Proceedings of Machine Learning Research Vol. 133 (PMLR, 2021), pp. 3–26.
- [74] A. Quitadadmo, J. Johnson, and X. Shi, Bayesian hyperparameter optimization for machine learning based eQTL analysis, in *Proceedings of the 8th ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics*, ACM-BCB’17 (ACM Press, New York, 2017), pp. 98–106.
- [75] B. J. Shields, J. Stevens, J. Li, M. Parasram, F. Damani, J. I. M. Alvarado, J. M. Janey, R. P. Adams, and A. G. Doyle, Bayesian reaction optimization as a tool for chemical synthesis, *Nature* **590**, 89 (2021).
- [76] P. M. Attia, A. Grover, N. Jin, K. A. Severson, T. M. Markov, Y.-H. Liao, M. H. Chen, B. Cheong, N. Perkins, Z. Yang, P. K. Herring, M. Aykol, S. J. Harris, R. D. Braatz, S. Ermon, and W. C. Chueh, Closed-loop optimization of fast-charging protocols for batteries with machine learning, *Nature* **578**, 397 (2020).
- [77] S. Ju, T. Shiga, L. Feng, Z. Hou, K. Tsuda, and J. Shiomi, Designing Nanostructures for Phonon Transport via Bayesian Optimization, *Phys. Rev. X* **7**, 021024 (2017).
- [78] A. Seko, A. Togo, H. Hayashi, K. Tsuda, L. Chaput, and I. Tanaka, Prediction of Low-Thermal-Conductivity Compounds with First-Principles Anharmonic Lattice-Dynamics Calculations and Bayesian Optimization, *Phys. Rev. Lett.* **115**, 205901 (2015).
- [79] F. Leclercq, Bayesian optimization for likelihood-free cosmological inference, *Phys. Rev. D* **98**, 063511 (2018).
- [80] R. A. Vargas-Hernández, Y. Guan, D. H. Zhang, and R. V. Krems, Bayesian optimization for the inverse scattering problem in quantum reaction dynamics, *New J. Phys.* **21**, 022001 (2019).
- [81] R. Mukherjee, F. Sauvage, H. Xie, R. Löw, and F. Mintert, Preparation of ordered states in ultra-cold gases using Bayesian optimization, *New J. Phys.* **22**, 075001 (2020).

- [82] R. Mukherjee, H. Xie, and F. Mintert, Bayesian Optimal Control of Greenberger-Horne-Zeilinger States in Rydberg Lattices, *Phys. Rev. Lett.* **125**, 203603 (2020).
- [83] P. B. Wigley, P. J. Everitt, A. van den Hengel, J. W. Bastian, M. A. Sooriyabandara, G. D. McDonald, K. S. Hardman, C. D. Quinlivan, P. Manju, C. C. N. Kuhn, I. R. Petersen, A. N. Luiten, J. J. Hope, N. P. Robins, and M. R. Hush, Fast machine-learning online optimization of ultra-cold-atom experiments, *Sci. Rep.* **6**, 25890 (2016).
- [84] R. J. Shaloo, S. J. D. Dann, J.-N. Gruse, C. I. D. Underwood, A. F. Antoine, C. Arran, M. Backhouse, C. D. Baird, M. D. Balcazar, N. Bourgeois, J. A. Cardarelli, P. Hatfield, J. Kang, K. Krushelnick, S. P. D. Mangles, C. D. Murphy, N. Lu, J. Osterhoff, K. Pöder, P. P. Rajeev *et al.*, Automation and control of laser wakefield accelerators using Bayesian optimization, *Nat. Commun.* **11**, 6355 (2020).
- [85] J. Duris, D. Kennedy, A. Hanuka, J. Shtalenkova, A. Edelen, P. Baxevanis, A. Egger, T. Cope, M. McIntire, S. Ermon, and D. Ratner, Bayesian Optimization of a Free-Electron Laser, *Phys. Rev. Lett.* **124**, 124801 (2020).
- [86] T. Mazza, T. M. Baumann, R. Boll, A. De Fanis, P. Grychtol, M. Ilchen, J. Montaño, V. Music, Y. Ovcharenko, N. Rennhack, D. E. Rivas, A. Rörig, P. Schmidt, S. Usenko, P. Ziolkowski, D. La Civita, M. Vannoni, H. Sinn, B. Keitel, E. Plönjes *et al.*, The beam transport system for the Small Quantum Systems instrument at the European XFEL: Optical layout and first commissioning results, *J. Synch. Radiat.* **30**, 457 (2023).
- [87] T. Mazza, H. Zhang, M. Meyer, and DESY, Technical Design Report: Scientific Instrument SQS, Tech. Rep. (2012).
- [88] T. Tschentscher, C. Bressler, J. Grünert, A. Madsen, A. P. Mancuso, M. Meyer, A. Scherz, H. Sinn, and U. Zastra, Photon beam transport and scientific instruments at the european XFEL, *Appl. Sci.* **7**, 592 (2017).
- [89] W. Decking, S. Abeghyan, P. Abramian, A. Abramsky, A. Aguirre, C. Albrecht, P. Alou, M. Altarelli, P. Altmann, K. Amyan, V. Anashin, E. Apostolov, K. Appel, D. Auguste, V. Ayvazyan, S. Baark, F. Babies, N. Baboi, P. Bak, V. Balandin *et al.*, A MHz-repetition-rate hard X-ray free-electron laser driven by a superconducting linear accelerator, *Nat. Photon.* **14**, 391 (2020).
- [90] A. C. LaForge, S.-K. Son, D. Mishra, M. Ilchen, S. Duncanson, E. Eronen, E. Kukkk, S. Wirok-Stoletow, D. Kolbasova, P. Walter, R. Boll, A. De Fanis, M. Meyer, Y. Ovcharenko, D. E. Rivas, P. Schmidt, S. Usenko, R. Santra, and N. Berrah, Resonance-Enhanced Multiphoton Ionization in the X-Ray Regime, *Phys. Rev. Lett.* **127**, 213202 (2021).
- [91] B. F. Murphy, T. Osipov, Z. Jurek, L. Fang, S.-K. Son, M. Mucke, J. H. D. Eland, V. Zhaunerchyk, R. Feifel, L. Avaldi, P. Bolognesi, C. Bostedt, J. D. Bozek, J. Grilj, M. Guehr, L. J. Frasinski, J. Glowonia, D. T. Ha, K. Hoffmann, E. Kukkk *et al.*, Femtosecond X-ray-induced explosion of C<sub>60</sub> at extreme intensity, *Nat. Commun.* **5**, 4281 (2014).
- [92] B. Rudek, K. Toyota, L. Foucar, B. Erk, R. Boll, C. Bomme, J. Correa, S. Carron, S. Boutet, G. J. Williams, K. R. Ferguson, R. Alonso-Mori, J. E. Koglin, T. Gorkhover, M. Bucher, C. S. Lehmann, B. Krässig, S. H. Southworth, L. Young, C. Bostedt *et al.*, Relativistic and resonant effects in the ionization of heavy atoms by ultra-intense hard X-rays, *Nat. Commun.* **9**, 4200 (2018).
- [93] T. Jahnke, R. Guillemin, L. Inhester, S.-K. Son, G. Kastirke, M. Ilchen, J. Rist, D. Trabert, N. Melzer, N. Anders, T. Mazza, R. Boll, A. De Fanis, V. Music, T. Weber, M. Weller, S. Eckart, K. Fehre, S. Grundmann, A. Hartung *et al.*, Inner-Shell-Ionization-Induced Femtosecond Structural Dynamics of Water Molecules Imaged at an X-Ray Free-Electron Laser, *Phys. Rev. X* **11**, 041044 (2021).
- [94] M. Hoener, L. Fang, O. Kornilov, O. Gessner, S. T. Pratt, M. Gühr, E. P. Kanter, C. Blaga, C. Bostedt, J. D. Bozek, P. H. Bucksbaum, C. Buth, M. Chen, R. Coffee, J. Cryan, L. DiMauro, M. Glowonia, E. Hosler, E. Kukkk, S. R. Leone *et al.*, Ultraintense X-Ray Induced Ionization, Dissociation, and Frustrated Absorption in Molecular Nitrogen, *Phys. Rev. Lett.* **104**, 253002 (2010).
- [95] S.-K. Son, R. Boll, and R. Santra, Breakdown of frustrated absorption in x-ray sequential multiphoton ionization, *Phys. Rev. Res.* **2**, 023053 (2020).
- [96] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning* (MIT Press, Cambridge, MA, 2006).
- [97] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey *et al.*, SciPy 1.0: Fundamental Algorithms for scientific computing in python, *Nat. Methods* **17**, 261 (2020).
- [98] T. Maltezopoulos, F. Dietrich, W. Freund, U. F. Jastrow, A. Koch, J. Laksman, J. Liu, M. Planas, A. A. Sorokin, K. Tiedtke, and J. Grünert, Operation of X-ray gas monitors at the European XFEL, *J. Synch. Radiat.* **26**, 1045 (2019).
- [99] M. Dommach, M. Di Felice, B. Dickert, D. Finze, J. Eidam, N. Kohlstrunk, M. Neumann, F. Meyn, M. Petrich, B. Rio, H. Sinn, and R. Villanueva, The photon beamline vacuum system of the European XFEL, *J. Synch. Radiat.* **28**, 1229 (2021).
- [100] L. Budewig, S.-K. Son, and R. Santra, Theoretical investigation of orbital alignment of x-ray-ionized atoms in exotic electronic configurations, *Phys. Rev. A* **105**, 033111 (2022).
- [101] L. Budewig, S.-K. Son, and R. Santra, State-resolved ionization dynamics of a neon atom induced by x-ray free-electron-laser pulses, *Phys. Rev. A* **107**, 013102 (2023).
- [102] A. Rörig, S.-K. Son, T. Mazza, P. Schmidt, T. M. Baumann, B. Erk, M. Ilchen, J. Laksman, V. Music, S. Pathak, D. E. Rivas, D. Rolles, S. Serkez, S. Usenko, R. Santra, M. Meyer, and R. Boll, Multiple-core-hole resonance spectroscopy with ultraintense x-ray pulses, [arXiv:2303.07942](https://arxiv.org/abs/2303.07942).
- [103] The data supporting the findings of the presented study are available under <https://doi.org/10.22003/XFEL.EU-DATA-002318-00>.
- [104] G. H. Yao and S. I. Chu, Generalized pseudospectral methods with mappings for bound and resonance state problems, *Chem. Phys. Lett.* **204**, 381 (1993).
- [105] X. M. Tong and S. I. Chu, Theoretical study of multiple high-order harmonic generation by intense ultrashort pulsed laser fields: A new generalized pseudospectral time-dependent method, *Chem. Phys.* **217**, 119 (1997).
- [106] J. W. Cooper, Photoionization from outer atomic subshells. a model study, *Phys. Rev.* **128**, 681 (1962).
- [107] S. T. Manson and J. W. Cooper, Photo-ionization in the soft x-ray range: Z dependence in a central-potential model, *Phys. Rev.* **165**, 126 (1968).

- [108] N. Rohringer and R. Santra, X-ray nonlinear optical processes using a self-amplified spontaneous emission free-electron laser, *Phys. Rev. A* **76**, 033416 (2007).
- [109] M. G. Makris, P. Lambropoulos, and A. Mihelič, Theory of Multiphoton Multielectron Ionization of Xenon under Strong 93-eV Radiation, *Phys. Rev. Lett.* **102**, 033002 (2009).
- [110] A. G. Kochur, A. J. Dudenko, V. L. Sukhorukov, and I. D. Petrov, Direct Hartree-Fock calculation of multiple  $\text{Xe}^{i+}$  ion production through inner shell vacancy de-excitations, *J. Phys. B* **27**, 1709 (1994).
- [111] G. Doumy, C. Roedig, S.-K. Son, C. I. Blaga, A. D. DiChiara, R. Santra, N. Berrah, C. Bostedt, J. D. Bozek, P. H. Bucksbaum, J. P. Cryan, L. Fang, S. Ghimire, J. M. Glowina, M. Hoener, E. P. Kanter, B. Krässig, M. Kuebel, M. Messerschmidt, G. G. Paulus *et al.*, Nonlinear Atomic Response to Intense Ultrashort X Rays, *Phys. Rev. Lett.* **106**, 083002 (2011).